DOCUMENT RESUME

ED 446 348 EA 030 616

AUTHOR	Fowler, William J., Jr., Ed.
TITLE	Developments in School Finance, 1998. Fiscal Proceedings from the Annual State Data Conference (July 1998).
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-2000-302
ISBN	ISBN-0-16-050487-2
PUB DATE	2000-08-00
NOTE	130p.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827. U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Educational Economics; Educational Equity (Finance); *Educational Finance; Elementary Secondary Education; Finance Reform; *Financial Policy; Needs Assessment; School Funds

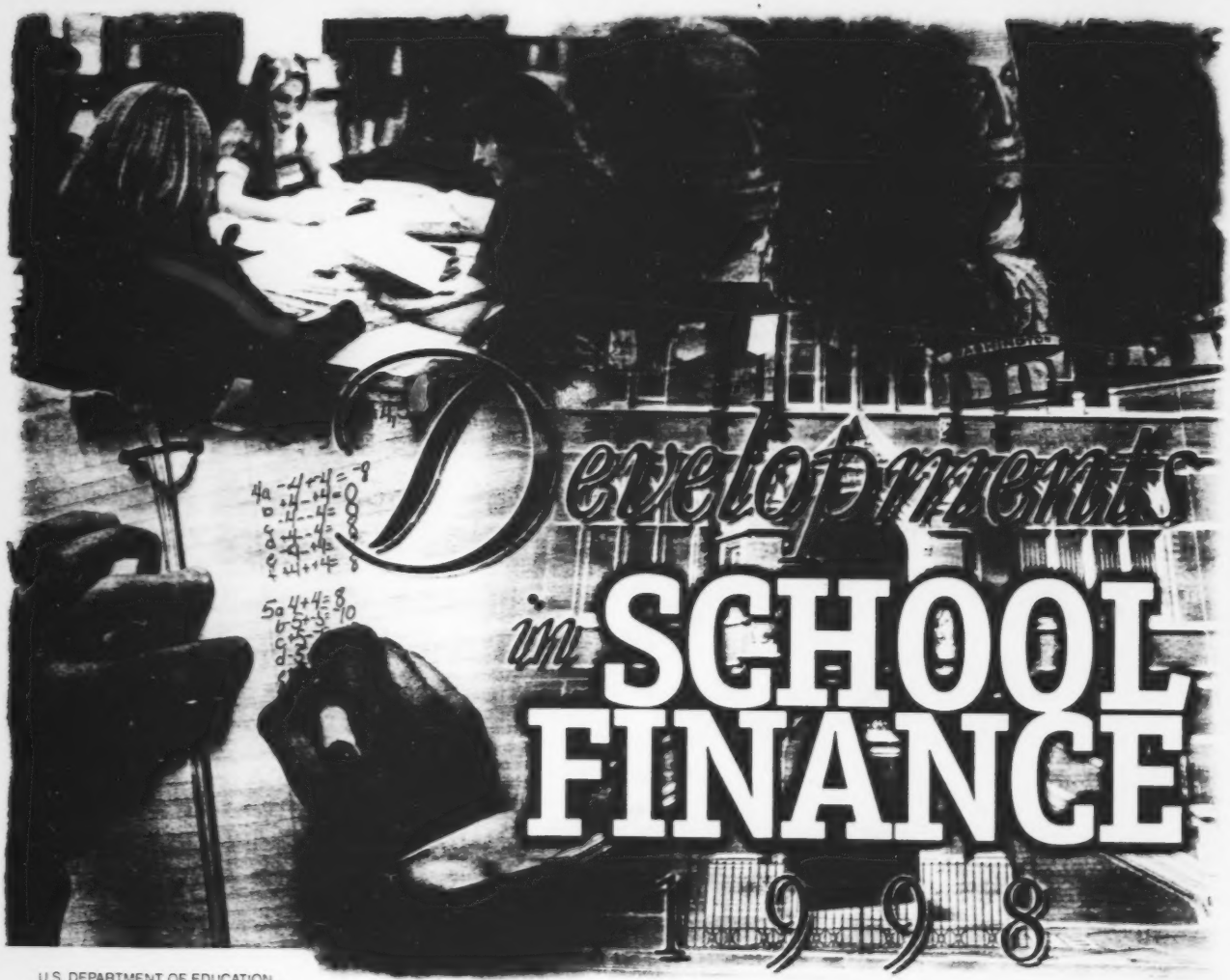
ABSTRACT

This collection of seven papers examines current theoretical perspectives on public-school finance. The first paper, "Reflections on the Limitations of Our Ability to Measure Schools' Productivity and Some Perspectives from the Past," assesses the inherent difficulties in measuring school productivity, particularly if one is concerned with more than test scores. The next paper, "Education Financing and Outcomes in Philadelphia," eschews efforts to measure productivity in the economic tradition, preferring to compare a school district with similar peer school districts in order to benchmark spending and performance. The third paper, "School Choice in Milwaukee: Are Private Schools Creaming Off the Best Students," discusses potential threats to the measurement of public school productivity, such as removing the most talented students through implementation of school-voucher programs. The following paper, "School-Level Resource Allocation in the Chicago Public Schools," focuses on a natural experiment in which funding responsibility was shifted from the district level to the school level. Two papers, "The Productivity of School Finance Equalization: An Analysis Using Hierarchical Linear Modeling," and "Active Graphics Methods for the Analysis and Display of Education Data," assess fiscal equalization. The final paper, "Developing Student Resource Variables for the Early Childhood Longitudinal Survey," describes how the National Center for Education Statistics has adopted a student-level finance measure that has the ability to accurately assess questions of the relationship of resources to student outcomes. (RJM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 446 348

National Center for Education Statistics



Developments in SCHOOL FINANCE 1998

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

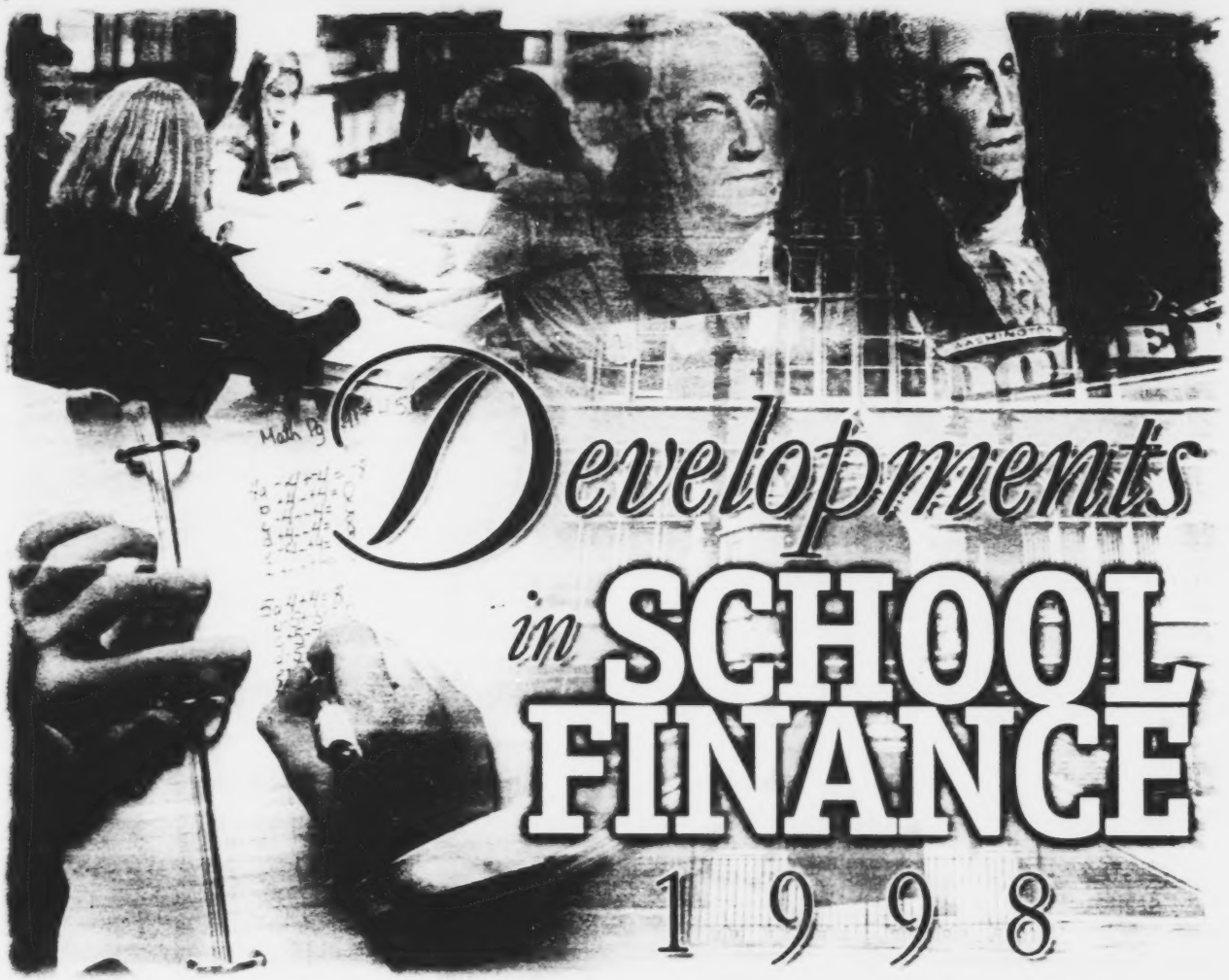
Fiscal Proceedings from the Annual State Data Conference July 1998

EA 030616

U.S. Department of Education
Office of Educational Research and Improvement
NCES 2000-302

BEST COPY AVAILABLE

National Center for Education Statistics



Developments in SCHOOL FINANCE 1998

William J. Fowler, Jr., Editor

Fiscal Proceedings from
the Annual State Data
Conference
July 1998

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2000–302

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
1990 K Street, NW
Washington, DC 20006–5651

August 2000

The NCES World Wide Web Home Page is: *http://nces.ed.gov*
The NCES education finance World Wide Web Home Page is: *http://nces.ed.gov/edfin*
The NCES World Wide Web Electronic Catalog is: *http://nces.ed.gov/pubsearch/index.asp*

Suggested Citation
U.S. Department of Education, National Center for Education Statistics. *Developments in School Finance, 1998*. NCES 2000–302. Ed. Fowler, William J. Jr. Washington, DC: 2000.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

or call toll free 1–877–4ED–Pubs

Content Contact:
William J. Fowler, Jr.
(202) 502–7338
Internet: William_Fowler@ed.gov

The papers in this publication were requested by the National Center for Education Statistics, U.S. Department of Education. They are intended to promote the exchange of ideas among researchers and policymakers. The views are those of the authors, and no official support by the U.S. Department of Education is intended or should be inferred.

BEST COPY AVAILABLE

Foreword

Jeffrey A. Owings, Associate Commissioner

Elementary/Secondary and Libraries Studies Division

Addressing the theme, "How to measure school performance in a tangible way," scholars in the field of education finance presented their thinking at the 1998 National Center for Education Statistics (NCES) Summer Conference. The implicit questions posed by all of the presentations revolve around the current and future financial status of school districts, how to portray that condition, and the significance of that standing for school performance.

Developments in School Finance contains papers presented at the annual NCES Summer Data Conference. This Conference attracts several state department of education policymakers, fiscal analysts, and fiscal data providers from each state, who are offered fiscal training sessions and updates on developments in the field of education finance. The presenters are experts in their respective fields, each of whom has a unique perspective or interesting quantitative or qualitative research regarding emerging issues in education finance. The reaction of those who attended the Conference was overwhelmingly positive. We hope that will be your response as well.

This proceedings is the fifth education finance publication from the NCES Summer Data Conference. The papers included within present the views of the authors, and are intended to promote the exchange of ideas among researchers and policymakers. No official support by the U.S. Department of Education or NCES is intended or should be inferred. Nevertheless, NCES would be pleased if the papers provoke discussions, replications, replies, and refutations in future Summer Conferences.

BEST COPY AVAILABLE

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328

ISBN 0-16-050487-2

BEST COPY AVAILABLE

Acknowledgments

The editor gratefully acknowledges the comments and suggestions of the reviewers: Leslie Scott of the Educational Statistical Services Institute (ESSI), and Michael P. Cohen who was employed by the National Center for Education Statistics (NCES) at the time of review. I also would like to acknowledge the contributions of the following individuals from Pinkerton Computer Consultants, Inc.: Janis Stennett wrote the introduction to the volume and assisted in editing the manuscript; Ross Pfile, Sonia Connor, and Rebecca Pratt edited the manuscript; Carol Rohr coordinated with each author to receive all materials and to resolve editorial issues; Allison Pinckney and Carol Rohr designed the layout, performed the desktop publishing, and prepared the files for printing; and Mark Ricks designed the cover.

Contents

Foreword iii

Acknowledgments v

Introduction and Overview 1
William J. Fowler, Jr.

Reflections on the Limitations of our Ability to Measure Schools' Productivity, and Some Perspective from the Past 9
Richard Rothstein

Education Financing and Outcomes in Philadelphia 19
Elizabeth Greenberg and John Guarnera

School Choice in Milwaukee: Are Private Schools Creaming Off the Best Students? 33
Dan D. Goldhaber, Dominic J. Brewer, and Eric R. Eide

School-Level Resource Allocation in the Chicago Public Schools 45
Ross Rubenstein

The Productivity of School Finance Equalization: An Analysis Using Hierarchical Linear Modeling 61
Patrick Galvin, Hal Robins, and Karen Callahan

Active Graphics Methods for the Analysis and Display of Education Data 81
Laurence Toenjes

Developing Student Resource Variables for the Early Childhood Longitudinal Survey 105
Lawrence O. Picus and Lauri Peternick

Introduction and Overview

William J. Fowler, Jr.

National Center for Education Statistics

About the Editor

William J. Fowler, Jr. is the acting program director of the Education Finance Program in the Elementary and Secondary Education and Library Studies Division at the National Center for Education Statistics (NCES), U.S. Department of Education. He specializes in elementary and secondary education finance and education productivity research. Dr. Fowler is currently involved with developing and implementing an individual student resource measure, and issuing a new NCES accounting handbook. In addition, he is engaged in devising a method of reporting education finance in a user-friendly language. His great passion is designing Internet tools for the NCES education finance site at http://nces.ed.gov/edfin, as well as the graphic display of quantitative data.

Dr. Fowler has worked at NCES since 1987. Prior to joining NCES, he served as a supervisor of school finance research for the New Jersey Department of Education. He has taught at Bucknell University and the University of Illinois, and served as a senior research associate for the Central Education Midwestern Regional Educational Laboratory (CEMREL) in Chicago and for the New York Department of Education. He received his doctorate in education from Columbia University in 1977.

Dr. Fowler received the Outstanding Service Award of the American Education Finance Association in 1997, and served on its Board of Directors from 1992 to 1995. He serves on the editorial board of the *Journal of Education Finance* and the *Educational Evaluation and Policy Analysis Journal* of the American Education Research Association. He also serves on the Board of Leaders of the Council for Excellence in Government, and was a 1997–98 Senior Fellow. He is a member of the Governmental Accounting Standards Board Advisory Committee which is charged with developing a *User Guide for Public School District Financial Statements*.

Introduction and Overview

William J. Fowler, Jr.

National Center for Education Statistics

The learned education finance researchers who presented at the 1998 National Center for Education Statistics (NCES) Summer Conference examined both current theoretical perspectives on public school finance and the everyday policy concerns of schools. Thus, the first four articles in this volume reflect policy studies of how to measure school performance in a tangible manner. The last three articles serve as theoretical explorations of cutting-edge research in the field of school finance. The implicit questions posed by all of these articles revolve around the current and future financial status for school districts, how to portray that condition, and the significance of that standing for school performance.

The first paper asserts how difficult it is to measure a school's productivity, particularly if one is concerned with more than test scores, such as constructive employment for graduates. Measuring school productivity is made more difficult because America's schools seek to achieve a wide array of inherently difficult-to-measure student outcomes, such as responsible citizenship. The next paper makes no effort to measure productivity in the economic tradition. Instead, it attempts to usefully compare a school district with similar peer school districts, in order to benchmark spending and performance. A third paper explores potential threats to the measurement of public school productivity, such as removing the most talented students through implementation of school voucher programs. Another scholar focuses attention on a natural experiment in Chicago's public schools, in which responsibility for the distribution of funding was shifted from the district level to the school level. This paper then examines school funding and its relationship to student achievement outcomes.

Two papers on the cutting edge of school finance research seek to better assess fiscal equalization. One employs hierarchical linear modeling (HLM), one of the most sophisticated statistical analytical tools available to educational researchers. HLM permits analysts to sort out differences that occur within and between organizational units, such as schools and school districts. The other paper explores methods of presenting financial data in graphic displays that permit widely diverse audiences, such as state legislators and parents, to quickly grasp meaning that can be used to improve school performance.

A final cutting-edge paper describes how the NCES Early Childhood Longitudinal Survey for Kindergarten (ECLS-K) has incorporated a student-level finance measure, which potentially has the ability to most accurately assess questions of the relationship of resources to student outcomes. In addition, observers in the sample of Kindergarten schools assess the adequacy of the school facilities, and their subjective assessment of the learning climate of the school. To education finance researchers, these data represent the pinnacle of the information pyramid, where information on students, parents, teachers, and schools can all be combined in a myriad of combinations to answer the multitude of questions in the field of education.

To summarize, the papers from this year's presentations sought to explore how to assess the productivity of schooling, whether through sophisticated econometric or statistical procedures, or simple comparison. Let us now turn to the specifics of each paper.

The publication begins with an overview of the difficulty in measuring a school's productivity. **Richard Rothstein**, of the Economic Policy Institute, asserts that the politicization of "educational productivity" has caused a mismatch in the way educational goals (inputs) are defined and the method by which results (outputs) are assessed. Although some products of an education system have concrete methods of assessment, such as test scores, it is a challenge to define many educational goals because the outputs of schools are, to some extent, inherently not measurable. Also, building a relationship between recognizable outputs and latent inputs is a complex process.

As a basis for this statement, Rothstein outlines the National Education Goals set forth by the National Education Goals Panel, which are a mix of concrete and abstract goals for the Nation's education system. Unfortunately, not all of these goals can be measured using traditional methods (e.g., proficiency test scores). More abstract goals, such as being prepared for responsible citizenship, are difficult to measure. It is important to recognize the effect these different types of goals have on each other, as they are not independent entities.

Rothstein argues that school finance scholars need to alter their focus from attaining more accurate calculations of school productivity to broadening how school productivity is gauged, and the type of data required. His examination of how "then and now" studies of education are conducted helps the reader place the satisfaction with the nation's public schools in historical perspective. These studies were initiated as a response to the often-heard public cry of "today's schools do not measure up to the standards of the past." Rothstein postulates that although these studies may have been methodologically unsound, chronicling these types of analyses provides meaning to the unchanging debate about the quality of American education. He concludes that, due to rapidly changing demographics, the increasing ineffectiveness of public school productivity studies over time emphasizes the need for the measurement of a broad range of school outputs and a more effective matching of school inputs and latent outputs.

Elizabeth Greenberg and **John Guarnera**, of American Institutes for Research, seek to make pragmatic benchmarking comparisons between comparable school districts with extant data. They investigate education financing and outcomes in Philadelphia and make comparisons among other big city school districts, other Pennsylvania school districts, and between cities and states.

Greenberg and Guarnera focus many of their comparisons on school-level characteristics, such as current expenditures, student/teacher ratios, student achievement, percentage of budget spent on instruction, dropout rates, and teacher absences. Furthermore, they examine the ratio of city school district administrative expenditures to instructional expenditures and include comparisons on demographic characteristics, such as population loss and poverty level.

Each comparison made provides the reader with greater insight into the complexity of the relationship of schools between or within cities and states. Comparing specific education indicators within a state is usually an easy task. However, while comparison among states is often desirable, it is difficult to make these comparisons because of the inconsistencies among the methods states use to calculate these indicators. Although states may measure spending differently, accurate comparisons of spending patterns can be calculated. For example, a city may be compared to its state's average, and then comparable city school districts compared after this ratio is calculated.

Greenberg and Guarnera demonstrate that such comparisons are difficult to perform, and require the application of much judgement. Although NCES has an Internet tool to compare school districts (see http://nces.ed.gov/edfin/search/search_intro.htm), Greenberg and Guarnera found they had to go far beyond this primitive web tool.

Dan D. Goldhaber, of the Urban Institute, **Dominic J. Brewer**, of RAND, and **Eric R. Eide**, of Brigham Young University, examine the issue of the creaming of academically talented public school students, encouraging them to move to private schools. As a setting for their research, Goldhaber, Brewer, and Eide examined the Milwaukee Parental Choice Program (MPCP). Opponents of choice programs often argue that such programs will exacerbate the flight of academically talented students from the public schools.

Prior research, which analyzed mathematics and reading test scores, has resulted in widely differing conclusions about the outcomes of participating in the Milwaukee choice program. Goldhaber, Brewer, and Eide believe that there are important differences in student characteristics between public and private schools. Some of these differences are easily observable and, therefore, easily measurable; however, there are unobservable differences that can cause biased estimates of the effects of private schooling. For example, parents with academically talented children might choose private sector schooling for their children more often than choosing to send their children to the Milwaukee public schools. Those parents who choose private sector schooling may be different in some subtle way from parents who choose public sector schooling. These differences are difficult to define and measure empirically, particularly if parental attributes are combined in datasets in some aggregate manner.

The design of the Milwaukee choice program may help researchers evaluate the existence of selection bias. To qualify for this program, students must meet poverty guidelines, not be enrolled in a private school in the immediate prior year, and be eligible to attend private, nonsectarian schools in the school district. As a result of the qualification criteria, the number of students admitted to this program is limited. If schools receive more applicants than their enrollment limit, they are required to accept students based on random selection. This random assignment of students permits researchers to compare students who applied and were not accepted with those who applied and were accepted with less concern about sample selection bias. Because all public school students who met the criteria for the program had the choice to participate, the Milwaukee program is especially useful for research into the issue of selection bias. However, Goldhaber, Brewer, and Eide conclude that the data yield is little indication that sample selection or creaming is associated with the decision to participate in the Milwaukee choice program.

Many in the school finance community recommend shifting responsibility for the distribution of funding from the district level to the school level. This decentralized system of allocation is being adopted in many schools throughout the country; however, there has not been much research into the methods by which school-level allocations are determined. **Ross Rubenstein**, of Georgia State University, examines the impact of such a shift on the distribution of funding to Chicago schools. Rubenstein's article delves into this issue of allocation and how differences between traditional district-level and newer school-level allocations affect performance equity among schools.

In the late 1980s, Chicago public schools were decentralized in order to shift the responsibilities for governance and improvement to the individual schools. This decentralized school system provides Rubenstein with a solid location for analyses of differences in expenditure patterns and student performance across schools.

When comparing general fund spending patterns with school performance, Rubenstein finds that the spending patterns among high and low performing schools are similar. Although small differences exist, a pattern does not emerge. This result is not surprising, for general fund allocations must be used to provide the same basic services to all students. A comparison of Chapter I spending reveals a more interesting pattern. Although both elementary and high schools spend the greatest proportion of their Chapter I funds on instruction, high schools distribute remaining funds evenly across all functional areas, whereas elementary schools tend to focus on instructional support and administration. Overall, when considering all types of funding and both elementary and secondary schools, Rubenstein finds a consistent pattern: Although there is not much variation in total spending among schools, there is variation in decisions concerning discretionary funds.

Patrick Galvin, Hal Robins, and Karen Callahan, of the University of Utah, examine the issue of school finance equalization by using hierarchical linear modeling (HLM). In their study, HLM is applied to control for differences in school district characteristics. Under the assumption that schools are hierarchically nested, two schools could appear exactly alike yet perform differently due to other contextual variables. The HLM method assists researchers by controlling for these contextual variables.

This article is based on the assumption that the underlying goal of school finance equalization is to promote educational achievement. Although there is concern about the fair distribution of resources, Galvin, Robins, and Callahan assert that the goal of equalization efforts is to equalize educational opportunities and outcomes. They propose that there is a mismatch in policy goals between

distribution of educational resources (equity) and concern about the use of these resources (productivity).

The authors find many advantages when using HLM to analyze school finance equalization. This method of analysis develops a relationship between distribution of inputs and an outcome measure. In addition, HLM permits the researcher to concurrently examine the effect of predictor variables on both the average intercept and slope of the relationship variable. The HLM model also permits exploration of interactions, and confounding effects of variables. For example, it is

> "...one thing to operate as a high-need school within a relatively wealthy environment and quite another to operate as a high-need school in a relatively poor environment."

Galvin, Robins, and Callahan seek to learn the ways in which schools, classrooms, and students function compared with their resource environment. By using the HLM model, the authors learn that the effects of resources can be hidden by interaction with other variables and that more extensive research into these interactions is needed in order to determine the relationship between educational resources and achievement. They find that one of the primary contributions of the use of HLM is that it changes the focus of the research from inequity in resources to the *relationship* between resources and outcomes.

The school finance research community often forgets that their audience is composed of policymakers, parents, and the general public. Because of the wide range of sophistication of these diverse audiences, research outcomes must be presented in concise, understandable ways. **Larry Toenjes**, of the University of Houston, in the final article of this publication, explores active graphics methods for the analysis and display of education data. Advances in and increased accessibility to personal computers have enabled education finance researchers to present data analyses in new, previously unimaginable ways. Advances in this technology are occurring so quickly that staying informed about the latest updates may overwhelm the user's ability to take advantage of what these new technologies offer.

Three of the most popular active graphics display methods are described in this article. The first method is linking, which allows a researcher to present two different relationships among two pairs of attributes of the data—how a change in one attribute would affect both relationships. The second method is brushing, which enables a researcher to control the position of an outline shape on the monitor screen so that the points descending within this shape can be visibly modified. Third, Toenjes highlights the spinning method, in which a researcher has the ability to rotate a three-dimensional figure around any or all of its axes.

Toenjes explores some of the software programs that employ these and other techniques. The history presented for these examples gives the reader a better understanding of the development and use of these programs. Following this background information, several sample applications and sample active graphics are presented to help the reader understand the benefits of real-life applications of these programs. For example, Toenjes displays the relationship in New York, Indiana, and New Jersey between the percent of local tax funds and the expenditure per pupil. The visual display clearly indicates that state aid has effectively equalized expenditures in New Jersey for poor school districts, but not in New York or Indiana. Using visual display, it is even possible to show specific school districts that are rich and poor, and contrast their characteristics, including per pupil expenditures. Although Toenjes concludes that these visual methods cannot replace the traditional techniques, they can facilitate understanding and communicating research results to a broad audience with wide differences in levels of statistical and research sophistication.

The final article in this publication brings the reader full circle. The publication began with a contemplation on limits to measure schools' productivity; it is concluded by **Lawrence O. Picus** and **Lauri Peternick's** suggestions for improving the Early Childhood Longitudinal Survey (ECLS) to enable improved measurement of the elusive variables many of the articles in this publication have described. Picus, of the University of Southern California, and Peternick, of American Institutes for Research, suggest the ECLS can be a valuable resource for collecting new finance data that will help uncover the relationship between student outcomes and resource allocation and use. By enhancing this existing survey with questions about how money matters in education, researchers would have the opportunity to delve into these latent

relationships of resource allocation and their effects on student achievement.

Picus and Peternick outline four broad categories for analysis. First, they propose examining classroom costs, which include teacher compensation, instructional aide compensation, instructional materials, and special programs. Second, they propose analyzing school-level costs, which include site administration, instructional support, student support, maintenance and operation, utilities, and transportation. Third, they propose analyzing district-level costs in terms of district administration, facilities, and data processing. Picus and Peternick also suggest analyzing nonschool costs, such as other agency expenditures and parental support. They conclude by giving us their recommendations for the items they think should be included in the ECLS.

Picus and Peternick then turn to the *actual* ECLS survey instruments containing fiscal data, that were devised by NCES after field-testing and comment by respondents. These survey instruments are mere shadows of the extensive items Picus and Peternick had recommended, but they avoid the excessive burden that would have been imposed on respondents by the many items they proposed.

As noted by Picus and Peternick, the collection and ultimate analysis of these data will provide a framework from which school districts can make wise resource allocation decisions, and in providing insight into understanding how and why resources matter.

Reflections on the Limitations of our Ability to Measure Schools' Productivity, and Some Perspective from the Past

Richard Rothstein

Economic Policy Institute

About the Author

Richard Rothstein is a research associate of the Economic Policy Institute, a senior correspondent of *The American Prospect*, and an adjunct professor of public policy at Occidental College in Los Angeles. Some material in the following article has been drawn from his book, *The Way We Were?*, published in 1998 by the Century Foundation, describing trends in measurement of student achievement over the last century. Rothstein's recent work on school expenditure trends was reported in Where's the Money Going?, a 1997 sequel to *Where's the Money Gone? Changes in the Level and Composition of Education Spending*, both published by the Economic Policy Institute.

Reflections on the Limitations of our Ability to Measure Schools' Productivity, and Some Perspective from the Past

Richard Rothstein

Economic Policy Institute

Claims about declining educational productivity are flawed for at least two reasons, each unremediable with present knowledge and data. First, we are quick to measure school outputs by math and reading scores, but require schools to produce a much broader range of outcomes, most of which are unmeasured and some of which are unmeasurable. Second, we have such poor longitudinal data on math and reading performance that claims based on declines in this performance rely more on selective anecdote than on representative data. Nonetheless, these unfounded claims have persisted for the better part of this century, with school observers in each era repeating the errors of the last.

The pages that follow attempt to illustrate each of these fatal limitations in our measurement of school productivity: first, the lack of definition and data relevant to many school goals; second, the persistent historical reliance on anecdote, not data, to support assertions of declines in the poorly measured outcomes of math and reading, and the attempts of school defenders to refute these anecdotal claims.

Even today, investigations of "educational productivity" are more primitive than many researchers and consumers of education research are inclined to acknowledge. The intense politicization of public education encourages making claims which surpass those supportable by unambiguous data. There are important public policy consequences to conclusions about whether schools have growing or declining productivity. Belief in the former encourages those seeking greater tax support for public schools. Belief in the latter encourages those seeking privatization or other radical structural reforms of schooling.

This heavy political burden borne by conclusions about school productivity is not a new phenomenon. With inadequate data, public education's critics have made claims regarding schools' "declining productivity" for nearly a century, and defenders have struggled to find evidence with which to refute them. It may be that the data available to us are too limited and imprecise to support legitimate conclusions about trends in education productivity. If this is the case, the contributions of quantifiable education research can play only a limited role in these debates. More important roles must be reserved for qualitative investigation and for the clarification of public values.

Trends in education productivity may be impossible to quantify with certainty because the outputs of schools

are mostly unmeasured. True, we devote great scholarly resources to measuring a few outputs such as standardized test scores in reading and mathematics, but there is much less interest in examining whether these outputs have a predictable relationship to schools' broad range of other outputs.

What are the outputs we seek from schools? The closest we have to an official list is the National Education Goals, first adopted by President Bush and the nation's governors in their 1989 Charlottesville meeting. Each year, the National Education Goals Panel produces a report on whether we are closer or farther from meeting these goals. The report essentially consists of "up" or "down" arrows to indicate whether or not progress is being made.

The following are the goals, as elaborated by the National Education Goals Panel:

- All children should start school ready to learn; they should have been born with an appropriate birthweight, should have access to high quality preschool programs, should have parents who have access to necessary training and support, and should receive the nutrition, physical activity, and health needed for mental alertness.
- Ninety percent of children should graduate from high school, and three-fourths of those who drop out will complete an equivalent degree; there will be no gap in high school graduation rates between minority and white students.
- All students should demonstrate competency over challenging subject matter (including reasoning, problem solving, and communication) in English, mathematics, science, foreign languages, civics and government, economics, arts, history and geography; they should be prepared for responsible citizenship (including participation in community service and activities demonstrating personal responsibility), further learning, and productive employment; achievement should improve for students in each quartile of the distribution, and the gap between performance of minority and white children should narrow; students should have access to physical and health education to ensure they are healthy and fit; more students should be competent in more than one language; and all students should be knowledgable about the diverse heritage of this nation and the world community.
- U.S. students will be first in the world in mathematics and science achievement; mathematics and science education will be strengthened, including use of the metric system of measurement; the number of qualified mathematics and science teachers will increase; the number of U.S. undergraduate and graduate students, especially women and minorities, who complete degrees in mathematics, science and engineering, will increase.
- Every adult will be literate, with skills needed to compete in a global economy and exercise his/her citizenship; connections between education and work will be strengthened, and quality public library programs for adults will grow in number. The proportion of students, especially minorities, entering and completing college will increase, and college graduates' critical thinking skills will improve. Schools themselves will offer more adult literacy and parent training programs.
- Every school will be free of drugs, violence, firearms, and alcohol and will teach drug and alcohol prevention as an integral part of health education. Schools will also eliminate sexual harassment.
- All teachers will have access to preservice teacher education and continuing professional development to provide the skills needed to teach a curriculum that can achieve the other goals.
- Every school will engage parents in a partnership that supports children's academic work and shared educational decision-making at school.

Even if we ignore items on this list that are not true outputs of schools (for example, whether children are born with appropriate birthweights), there are enough real outputs to make it evident that mathematics and reading scores alone cannot be the numerator in a measure of schools' productivity.

Could it be, however, that mathematics and reading scores are also proxies for other outputs, so we can presume that if mathematics and reading scores rise or decline, other outputs will also follow a similar trajectory? Not only is this not necessarily the case; the opposite may be true. Systems with complex combinations of goals must

always be wary of measuring only a few, because there will be powerful pressures on the system to focus on enhancing the production of the most easily measured outputs, at the expense of others that are equally important, but more difficult to measure. Everyone has heard stories of how "accountability for results" distorted enterprises in the old Soviet economy. When shoe factories were told they would be held accountable for the number of shoes they produced, the factories responded by producing only small sizes—gaining rewards for exceeding quotas without having to purchase more leather.

As the interest of both policymakers and scholars in school productivity outpaces our ability to measure it, American education faces similar dangers. Consider this disturbing set of facts: In the last decade, American secondary schools have come under great public and political pressure to increase the number of academic courses high school students must take to qualify for a diploma. This pressure has worked. From 1988 to 1994, the percentage of school districts requiring at least 4 years of secondary school English for a high school diploma rose from 80 to 85 percent; those requiring at least 3 years of mathematics rose from 35 to 45 percent; and those requiring at least 3 years of science rose from 17 to 25 percent (U.S. Department of Education 1998, Indicator 26). Most Americans consider these data a sign of progress, and considered by themselves, they certainly are. But our national goals tell us that we also want students to have "access to physical and health education to ensure they are healthy and fit." Nearly simultaneously with the increase in academic course-taking, the percentage of students taking a daily high school physical education course declined from 42 to 25, and the number of overweight adolescents soared (Sammann 1998, 2). We do not know why this is the case, but one possibility is that pressure to require more mathematics and science courses to improve mathematics and science scores has led American high schools to find the time for these courses by reducing requirements for physical education. This result may not be consistent with Americans' goals for education, which include *both* mathematics and science proficiency at ever higher levels, *and* the development of habits that lead to lifelong good health. If we hold schools accountable only for the former, and find it relatively easier to measure these, we run the risk of de-emphasizing other less important aspects of achievement far more than a balanced program would require, no matter how important mathematics and science proficiency may be.

In the last decade, American secondary schools have come under great public and political pressure to increase the number of academic courses high school students must take to qualify for a diploma.

In truth, we have devoted little energy to attempts to measure many of the important outputs of schools. Not only do we have no standardized reports of adolescents' physical health, but we have no (or very limited) trend data on the national goals of "responsible citizenship," avoidance of drug, alcohol, and tobacco abuse, competency in fields like the arts, or "knowledge about the diverse heritage of this nation and the world community," or "participation in community service and activities demonstrating personal responsibility." The National Assessment of Educational Progress (NAEP) has begun to test a few of these other areas (the new arts and music assessment is one example), but few states have high stakes tests that go beyond reading and mathematics, so the potential for distortion of school production in favor of those outputs most easily measured is real.

Conclusions about school productivity require not only measurement of outputs but measurement of inputs. As I have argued elsewhere, because we have limited ability to match inputs with the particular outputs they are designed to produce, school productivity becomes even a more elusive concept. We might make some tentative conclusions about school productivity if we tracked mathematics and reading scores and compared them with changes in the schools devoted to enhancing mathematics and reading achievement, but data on resources, reported by function and object, not by program, are ill-suited for this purpose (Rothstein and Miles 1995).

The aim of this paper is to urge school finance scholars to ease off, a little, in the quest for more precise calculations of school productivity. Before making existing data more precise, more energy should be invested in broadening the data we require.

In what follows, I attempt to place the desire for conclusions about school productivity in historical perspective.

BEST COPY AVAILABLE

We are not the first generation of education researchers to confront the need for data to respond to a broad public concern that school productivity is declining. School administrators and researchers frequently attempted to analyze the persistent claims of declining student performance by conducting "then and now" studies of education.

In reading through the educational research literature of the past, one cannot help but be struck by the consistently defensive tone of many "then and now" studies: professors or school district administrators began their published reports by recounting the denunciations of schools by politicians, journalists, academics, pundits, or parent groups who had claimed that "today's" schools do not measure up to the standards of the past, that teachers no longer instructed children in basic skills, and that young people knew less than they once did. The professors or administrators then exclaimed that they had been subjected to this abuse long enough, and therefore had combed school district archives for tests given to students several decades earlier. The researchers then described how they had administered these outdated tests to contemporary students, under conditions as similar to those of the past as possible. The reports frequently concluded by showing the contemporary scores to be superior, refuting the conventional wisdom of the day.

While most, ... of these "then and now" studies were methodologically unsophisticated, ... the history of these studies sheds a useful light on the unchanging debates about the quality of American education.

While most, though not all, of these "then and now" studies were methodologically unsophisticated, and while careful social scientists today would never sanction such "uncontrolled" (for background characteristics) research, the history of these studies sheds a useful light on the unchanging debates about the quality of American education. Also, if we make the not-unreasonable assumption that demographic change may not have been as rapid during the period of these earlier studies as it is today, the studies, on the whole, suggested that school critics of the past were mistaken.

There have been several dozen such investigations; I will here describe only a few typical examples. The baseline was established by America's very first standardized test, administered in 1845 to a select group of 500 of Boston's brightest 8th-graders. Results were disappointing. The testing committee reported it "difficult to believe that there should be so many children ...unable to answer, ...so many absurd answers, so many errors in spelling, in grammar, and in punctuation." Only 45 percent of these top 14-year-olds knew that water expands when it freezes. When children did answer a question correctly, they frequently did not understand the answer they had given, because, as the examining committee put it, the children had been taught "the name of the thing rather than the nature of the thing." According to Massachusetts Secretary of Public Instruction Horace Mann, Boston's schools were ignoring higher-order thinking skills; what little students knew came from memorizing "words of the textbook, ...without having ...to think about the meaning of what they have learned." Thus 35 percent knew from history classes that, prior to the War of 1812, the United States had imposed an embargo on British and French shipping, but few had any idea what "embargo" meant. In one school, 75 percent of the students knew the date of the embargo, but only 5 percent could define the term (Caldwell and Courtis 1924, 52, 54, 90, 125).

In 1924, Otis Caldwell, a school director and Columbia Teachers College professor in New York City, and Stuart Courtis, director of teacher training at Detroit Teachers College, noted that "[s]urvey after survey has revealed unsuspected inadequacy or inefficiency in American education," resulting in "[s]uperintendents and teachers [being] dismissed" and "school systems and methods [being] reorganized." Caldwell and Courtis determined to "bring a long-delayed message of encouragement to all who have participated in accomplishing the educational progress of the last fifty years." To do so, they uncovered the test given by Horace Mann's committee of examiners in 1845, and re-administered this 1845 Boston test to a national sample ("from Maine to California") of 8th-graders in 1919 (Caldwell and Courtis 1924, v, vi, 8, 9, 77).

Mann's test questions that had retained curricular relevance 75 years later were selected—questions like those asking students to describe the "height of a heavenly body," or, "how high can you raise water in a common pump, with a single box?," or about the invasion of Canada in the "last war" were dropped. Caldwell and

BEST COPY AVAILABLE

Courtis printed a new exam with these remaining questions, and invited school districts to participate. School superintendents from 46 states volunteered. Unlike the Mann test, which had been given only to the best "brag scholars" (students whom Mann described as "the flower of the Boston schools"), the superintendents agreed to administer the test to all 8th-graders who were present on the day the test was given. Twelve thousand exams were returned for scoring.

Caldwell and Courtis found that despite the fact that the test in 1919 had been administered to a full range of 8th-graders, not only the brightest as in 1845, the median score on the still-relevant questions had been 37.5 percent correct in 1845, but 45.5 percent in 1919. They concluded that children in 1919 did somewhat worse than the earlier children in "pure memory" questions, and somewhat better on the "thought or meaningful questions." With respect to my earlier example, the researchers reported that "in 1845, 35 percent of the children knew the year when the embargo was laid by President Jefferson, but only 28 percent knew what an embargo was. In 1919, only 23 percent knew the year, but 34 percent knew the meaning" (Caldwell and Courtis 1924, 85–87).

Tests of General Educational Development (GED) are used as an alternative high school certification for students who drop out..

In 1934, a Los Angeles school researcher, Elizabeth Woods, gave a 1924 6th-grade reading test to students in 33 elementary schools in which the test had been administered ten years earlier. She found that scores were half a grade higher in 1934 than they had been in 1924 (Raths and Rothman 1952; Gray and Iverson 1954).

In 1946, Don Rogers, a Chicago assistant school superintendent, tired of hearing "employers... allege that present-day pupils (even high school graduates) are not proficient... The imputation is that ...our school system formerly trained them better than now" (Rogers 1946). So Rogers re-administered a 6th-grade Chicago arithmetic test from 1923. He found that the 1946 pupils on average scored about the same as 1923 pupils (despite the unusually high teacher turnover the 1946 students had experienced during World War II, and the constant disruptions of wastepaper, soap, and other wartime drives conducted in schools) and concluded that this test "discounts the allegations that ...Chicago pupils of an earlier generation did better work than their sons and daughters who are now in the elementary schools."

In 1948, Springfield (Missouri) schools came under attack from a citizens group for embracing tenets of "progressive education" and for ignoring the teaching of basic skills, particularly in reading. University of Illinois Professors F. H. Finch and V. W. Gillenwater undertook a study to "reveal whether the teaching of reading had increased or decreased in effectiveness," by re-administering a 1931 6th-grade reading test to contemporary 6th-graders in the same Springfield schools. They found that 1948 students had higher scores, and concluded that "[a]pparently reading instruction ...is now more effective... and most sixth grade children now in schools do better in reading than did their predecessors" (Finch and Gillenwater 1949). While Finch and Gillenwater did not use formal statistical controls that we would expect in such research today, they superficially investigated the characteristics of 1931 and 1948 students, and determined that the occupational classifications of the parents were similar in the two years.

Tests of General Educational Development (GED) are used as an alternative high school certification for students who drop out. They were originally developed in 1943 by the Army to assess the academic skills of draftees. To establish a scoring scale, the Army contracted for the test to be administered to a representative sample of 35,000 seniors in 814 high schools across the country (representative, that is, except that in segregated states only white schools were included). In 1955, at a time of ferocious public criticism of the public schools (and when the belief that schools had deteriorated was as widespread as it is today), Army officials wondered if the 1943 scale was still appropriate. So the Department of Defense contracted with the University of Chicago to conduct a new study, giving a 1955 GED test to a similarly representative national group of seniors. Then, a smaller sample of students were given both the 1943 and 1955 tests, so that the scales on the two could be equated. The Chicago professor who analyzed the results, Benjamin Bloom, concluded that "[I]n each of the GED tests the performance of the 1955 sample of Seniors is higher

BEST COPY AVAILABLE

than the performance of the 1943 sample... [I]n mathematics the average senior tested in 1955 exceeds 58 percent of the students tested in 1943." (Average performance also exceeded earlier scores in natural sciences, literary materials, English, and social studies.) "These differences are not attributable to chance variation in test results," Bloom concluded, and "indicate that the high schools are doing a significantly better job of education in 1955 than they were doing in 1943" (Bloom 1956).

In the early 1950s, Vera Miller and Wendell Lanton, researchers working for the Evanston (Illinois) school district, noted that parents and educators often charged that "too much time [is] being devoted to music, arts, crafts, dramatics and unit work [group projects] to the detriment of the 'Three R's.'" In response, Miller and Lanton reprinted the standardized reading tests that had been given 20 years earlier in the district, and re-administered them to contemporary students. Like Finch and Gillenwater in Missouri several years earlier, they also had no formal statistical controls for background characteristics, but they noted that the "community was relatively stable [and] present day groups of pupils and those of the past were similar in most respects... The area contains a cross section of people of different races and of varied social and economic status." To assure the most practically consistent test conditions, the district's testing director from the 1930s also administered the test in the 1950s, using similar procedures. The 1950s tests were given on or near the same day of the month as the 1930s tests. Miller and Lanton tested 3rd-, 4th- and 8th-graders from 1952 to 1954, and found that, for example, 4th-graders in 1952 scored 6 months higher in reading comprehension, and 8 months higher in vocabulary, than did their 1932 counterparts. "[P]resent day pupils read with more comprehension and understand the meaning of words better than did children who were enrolled in the same grades and schools more than two decades ago," Miller and Lanton concluded (Miller and Lanton 1956).

In 1976, the Indiana state Superintendent of Public Instruction, Harold Negley, teamed with two Indiana University professors, Roger Farr and Leo Fay, to examine the state's reading instruction. Their report notes that in 1976, "the charge is sometimes made that children do not read as well as in the past and that schools are to blame" (Farr and Negley 1978). In 1945, the state had administered a standardized reading test to a sample of 25 percent of the state's students at each grade level. In 1976, Farr, Fay, and Negley reprinted the 1945 tests for the 6th- and 10th-grades, and re-administered the tests to a comparable sample of students. The new sample, which included 7 percent of the state's students in those grades, was selected to be representative of the state's regional diversity and urban-rural-suburban distribution. Thus the demographic of the test-takers in the two time periods was as similar as possible. Their raw results revealed that 6th- and 10th-graders in 1976 read at virtually the same grade level as comparable students in 1945. For example, the average 1945 6th-grader read at exactly the national 6th-grade norm that had been established in 1943, while the average 1976 6th-grader read at one-tenth of one month below the 1943 6th-grade norm.

"[T]he charge is sometimes made that children do not read as well as in the past and that schools are to blame."

The state of Indiana, however, had kept unusually good records on the students who took the 1945 test, and Farr, Fay, and Negley noted that these students were considerably older than the 6th- and 10th-grade students who took the test in 1976. In 1945, it was more common not to promote students whose achievement was below grade level than it was in 1976: in the latter year, for example, the 6th-grade included 11- and 12-year-olds almost exclusively, but in 1945 there had been many 13- and 14-year-olds in the 6th-grade as well. In the 1940 census, average Indiana 6th-graders were 12 years and 4 months old, but in the 1970 census, they were only 11 years and 6 months old, nearly a full year's difference in average age. Consequently, the older 1945 students had been in school more years than the "comparable" 1976 students. Further, because fewer students dropped out between 9th- and 10th-grade in the later than in the earlier year, the 1945 10th-grade students were, on average, higher achievers, relative to all young people their age, than were the less selective group of 1976 10th-grade students. When Farr, Fay, and Negley adjusted their results to compare "age equivalent" scores rather than "grade equivalent" scores, they found that the 1976 sample, for both 6th- and 10th-grade, "outscored the 1945 sample significantly on every test."

BEST COPY AVAILABLE

"[T]he general national assumption that the reading abilities of our children are decreasing at an alarming rate [is] unsupported by this study," the Indiana researchers concluded. This "ungrounded alarm," however, "leads to attacks on school programs that have been developed over the same time span for which this study shows the improvement in student reading achievement."

Over the years, a few "then and now" studies have shown declining student achievement: a St. Louis Board of Education study, for example, found that reading achievement was slightly less in 1938 than it had been in 1916 (Boss 1940). However, most of these reports claimed improvement, to refute widely publicized attacks on schools in each era.

School officials and researchers no longer publish such reports, perhaps because we recognize that in order to make reasonable "then and now" comparisons of test scores, we require more sophisticated controls than the informal demographic similarities noted in the earlier studies. Especially because demographic change in many districts and schools has been more rapid in recent years, we can no longer take "then and now" studies seriously without better data on test takers' parental education, occupation, and even income, as well as the children's race and ethnicity, family status, and other socio-economic characteristics. These data simply do not exist for past test scores, and there is no way to create them.

Nor was it the case that these old "then and now" studies were conducted when it was possible to state without equivocation that school productivity could be measured solely in the fields of mathematics and reading. Many of these studies were conducted during the height of progressive education's influence, when the "Americanization" of students and the delivery of a broad range of social services through the schools were considered central to their mission.

Nonetheless, then, as now, there was a public demand for higher school standards in reading and mathematics (referred to as the "Three R's"), and the education policy community conducted its debates as though reading and mathematics were the only goals for which schools could be held accountable. Then, as now, the debates were ultimately unsatisfactory. They will continue to be unsatisfactory, until we can measure the broad range of school outputs, and match disaggregated inputs to the outputs they are designed to achieve.

References

Bloom, B. S. March 1956. "The 1955 Normative Study of the Tests of General Educational Development." *The School Review* 64.

Boss, Mable E. January 13, 1940. "Reading, Then and Now." *School and Society.*

Caldwell, Otis W. and Stuart A. Courtis. 1924. *Then and Now in Education*, 1845–1923. Yonkers-on-Hudson, New York: World Book Co.

Farr, Roger, Leo Fay, and Harold Negley. 1978. *Then and Now: Reading Achievement in Indiana* (1944–45 and 1976). Bloomington, IN: School of Education, Indiana University.

Finch, F. H. and V. W. Gillenwater. April 1949. "Reading Achievement Then and Now." *Elementary School Journal*, 49(8):446–454.

Gray, William S. and William J. Iverson. 1954. "How Well Do Pupils Read?" In C. Winfield Scott and Clyde M. Hill, eds., *Public Education Under Criticism*. Englewood Cliffs, NJ: Prentice Hall.

Miller, Vera V. and Wendell C. Lanton. February 1956. "Reading Achievement of School Children—Then and Now." *Elementary English*, 33.

Raths, Louis E. and Philip Rothman. April 1952. "Then and Now." *NEA Journal.*

Rogers, Don C. October 1946. "Success or Failure in School." *American School Board Journal.*

Rothstein, Richard, with Karen Hawley Miles. 1995. *Where's the Money Gone? Changes in the Level and Composition of Education Spending.* Washington, DC: Economic Policy Institute.

Sammann, Patricia. 1998. *Active Youth. Ideas for Implementing CDC Physical Activity Promotion Guidelines.* Champaign, IL: Human Kinetics.

U.S. Department of Education, National Center for Education Statistics. *The Condition of Education 1998.* NCES 98–013. Washington, DC: 1988.

Education Financing and Outcomes in Philadelphia

Elizabeth Greenberg

John Guarnera

American Institutes for Research

About the Authors

Elizabeth Greenberg is a senior research analyst at the American Institutes for Research (AIR), where she has been employed since 1997. She specializes in designing surveys and performing data analysis. Prior to her employment at AIR, she worked for the Economic Research Service of the Department of Agriculture studying regional economic and educational issues. Her recent publications have focused on the analysis of various National Center for Education Statistics data sets, including the National Adult Literacy Survey and the National Assessment of Educational Progress.

John Guarnera is a Credit Risk Analyst for First Union National Bank where he specializes in the monitoring of the credit risk associated with derivatives products. He formerly worked as a Research Assistant at the American Institutes for Research (AIR) where he specialized in the analysis of education finance issues. His most recent publication focused on school-level data collection efforts in the state of Ohio.

BEST COPY AVAILABLE

Education Financing and Outcomes in Philadelphia

Elizabeth Greenberg

John Guarnera

American Institutes for Research

Introduction

In the spring of 1998, the state of Pennsylvania passed legislation giving it the power to take over the city of Philadelphia's public schools because of poor student performance. This paper was written in response to that legislation. It attempts to assess the performance of Philadelphia's public schools using easily available public data, in order to determine if the concerns of the state of Pennsylvania were justified.

The first part of the paper compares Philadelphia's public schools with schools in other large northeastern cities. The cities in the comparison data set were chosen because they share with Philadelphia two demographic characteristics that are indicative of cities in trouble: high poverty rates and recent population loss. Philadelphia is compared with these cities in terms of current expenditures per pupil, student/teacher ratio, and average SAT scores.

None of the cities in this comparison group is in the same state as Philadelphia. Since the United States does not have a uniform group of assessments taken by students in different states, it is difficult to compare the achievement of Philadelphia's students with the achievement levels of students in these other cities. SATs are taken by high school juniors and seniors across the country. However, all students do not participate in the SATs; thus comparisons based on these tests may be misleading.

Therefore, an additional set of school districts within Pennsylvania, whose students were assessed using the same tests as the students in Philadelphia, was chosen as a second comparison group for Philadelphia. This group included the second and third largest cities in Pennsylvania, Pittsburgh and Harrisburg, since large cities often face similar problems in their educational systems. It also included two of the largest Philadelphia suburban school districts, because some educational issues such as the supply of qualified teachers vary by geographic area. Finally, it included the suburban Philadelphia public school district, Chester-Upland, with poverty rates closest to Philadelphia, as many studies have shown that poverty levels are one of the demographic characteristics most correlated with student achievement.

The analysis in this paper does not pretend to be a definitive assessment of the quality of education in Philadelphia. Such an assessment would require more time than was available to write this paper, and would also require access to data that is not in the public domain. The analysis in this paper is only intended to determine

BEST COPY AVAILABLE

if Philadelphia appears to be so outside the realm of expected performance that emergency measures are necessary.

However, this paper does illustrate a methodology that can be applied quickly to any school district to determine if its performance is so bad when compared with other districts that outside intervention may be necessary. Any analysis such as this should be followed up by more careful study including an examination of curriculum, facilities, staff qualifications, and other factors.

Comparisons between Philadelphia and Other Big City School Districts

Population Loss

In the fall of 1994, 208,000 students were enrolled in the Philadelphia School District, placing it among the top 10 school districts in the United States in terms of size. Although Philadelphia is still large, the city has been losing population. Between 1980 and 1992, the population of Philadelphia declined by 8 percent. As illustrated in figure 1, this drop in population also occurred in other large northern and eastern cities. Because cities that are losing population face a declining tax base and similar fiscal constraints, we chose this group of population loss cities as the comparison set for Philadelphia in this study.

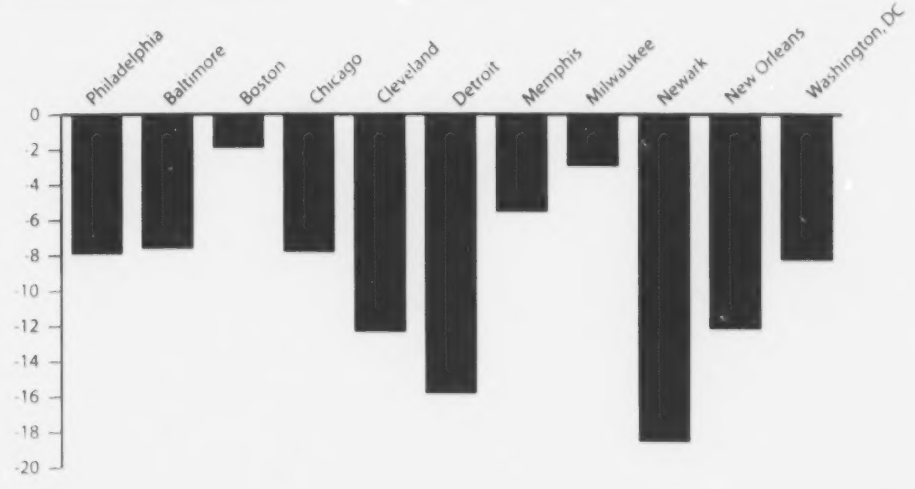
Poverty Level

In addition to the fact that they are all losing population, the set of cities is also characterized by the high percentage of children living below the poverty line. Thirty percent of children in Philadelphia live in homes with incomes below the poverty line, compared with a low of 25 percent and a high of 46 percent among other cities in the comparison group (figure 2).

Current Expenditures

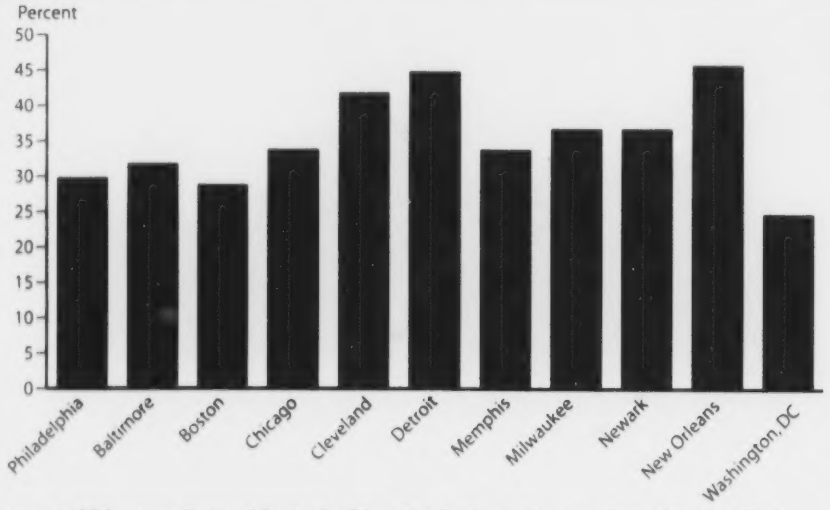
Philadelphia's current expenditures per pupil are in the middle of the group of comparison school districts (figure 3). In 1992–93, among school districts in the comparison group, Newark, Washington, DC, Boston, and Milwaukee spent more per pupil than Philadelphia.

Figure 1.—Percentage change in population, 1980–92



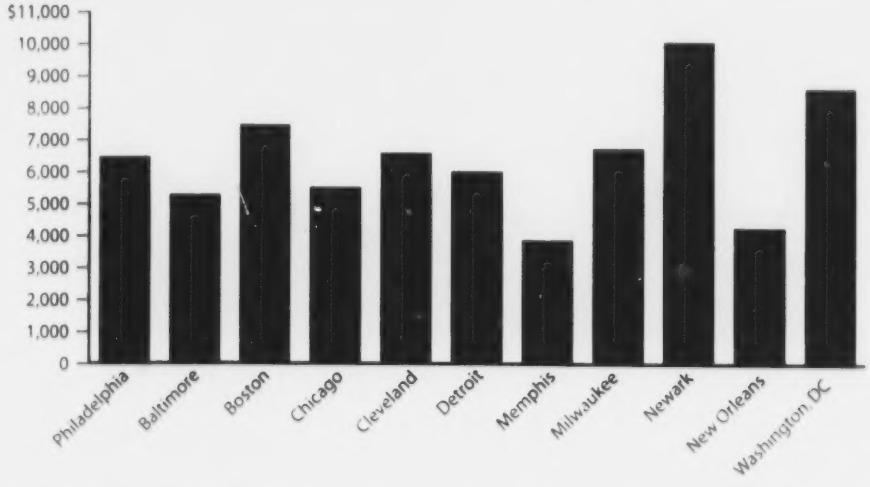
SOURCE: City and County Data Book, 1994.

Figure 2.—Percentage of children below poverty level



SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1992–93.

Figure 3.—Current expenditures per pupil



SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1992–93.

Student/Teacher Ratio

However, despite Philadelphia's relatively high current expenditures per pupil, among the comparison cities only Memphis had a higher student:teacher ratio (figure 4). In 1992–93, the most recent year for which we have been able to obtain data, the Philadelphia Public Schools had an average of 18.5 students per teacher. As illustrated in figure 3, Cleveland and Milwaukee had almost identical expenditures per pupil during that year, but Cleveland had a ratio of only 14.4 students per teacher and Milwaukee had a ratio of 16.6 students per teacher (figure 4).

Student Achievement

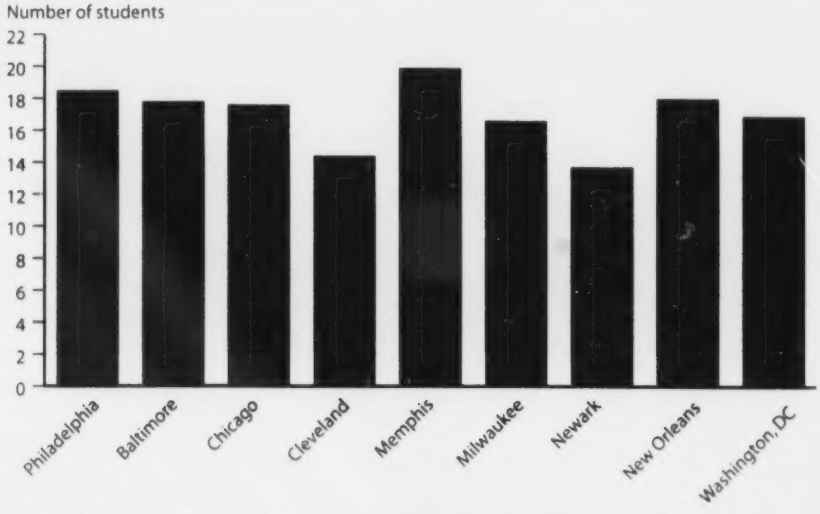
The Philadelphia Public School District and the other school districts in the comparison group use different standardized tests to assess their students; therefore it is difficult to compare academic outcomes among the school districts. However, students living in most of these school districts who intend to go to college take the SAT examination. While SAT scores are not the best comparison across school districts because the population taking them is self-selected and may vary significantly across the district, we can safely say that Philadelphia's average composite SAT scores are in line with those of other districts in the comparison group (figure 5).[1]

Comparisons between Philadelphia and Other Pennsylvania School Districts

Choosing within State Schools

Because much of the data collected by governments in the United States is collected at the state level, it is easier to compare the Philadelphia Public School District with other school districts in the state of Pennsylvania than it is to compare it with large city school districts in other states. As our comparison group of school districts within Pennsylvania, we picked two other city school districts, Pittsburgh and Harrisburg, as well as three suburban Philadelphia school districts. Pittsburgh and Harrisburg were selected because they are the second and third largest cities in Pennsylvania, after Philadelphia. Two of the suburban districts, Cheltenham and Abington, are primarily middle class, although they include pockets of low-income families. They were chosen because they were among the largest districts geographically close to Philadelphia. One of the suburban districts, Chester-Upland, is a poor district and was chosen because its poverty rates were similar to Philadelphia's.

Figure 4.—Student/teacher ratio



SOURCE: U.S. Department of Education, National Center for Education Statistics, Common Core of Data, 1992–93.

[1] Data on SAT scores is from 1995, while data on expenditures, student/teacher ratio, and poverty levels is from 1992–93. We used data for the most recent years available. Students' SAT scores are influenced by their entire educational and life experience, not just their experience in the current school year.

Poverty Level

Figure 6 illustrates the percentage of low-income students in each of the Pennsylvania school districts in our comparison group.

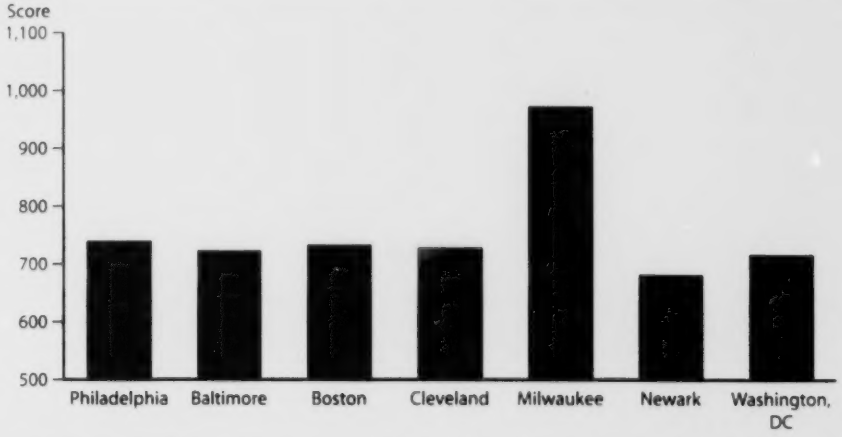
Current Expenditures

As illustrated in figure 7, Philadelphia's current expenditures per pupil during 1995–96 were lower than any of the other Pennsylvania school districts in the comparison group. Philadelphia spent $6,550 per pupil during 1995–96. Of the other districts in the Pennsylvania comparison group, only Chester-Upland spent less than $7,000 per pupil. Pittsburgh, the second largest city in Pennsylvania, spent $9,500 in 1995–96, almost one-third more than was spent by Philadelphia.

Percentage of Budget Spent on Instruction

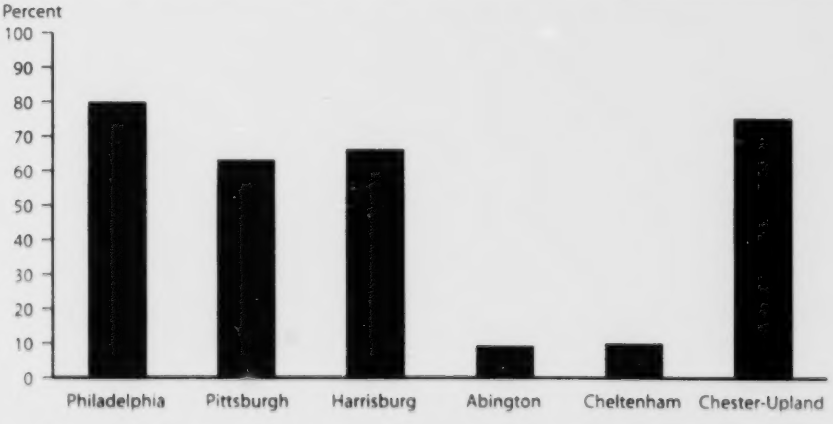
Although Philadelphia spent less per pupil in 1995–96 than the other school districts in the Pennsylvania comparison group, the percentage of its budget that Phila-

Figure 5.—Average composite SAT scores, 1995



SOURCE: Dr. Joyce Ladner, "Financing Education in the District of Columbia from the Perspective of the Financial Authority," presentation before the American Education Finance Association Annual Conference, March 7, 1997.

Figure 6.—Percentage of students living in low-income homes in selected Pennsylvania districts, 1996–97



SOURCE: Pennsylvania System of School Assessment School Profiles 1996–97.

delphia spent on instructional activities was about average for this group of school districts. In 1995–96, Philadelphia allocated 65 percent of its expenditures to instructional activities, compared with a high of 70 percent for Harrisburg and a low of 62 percent for Pittsburgh (figure 8).

Drop-out Rates

Figure 9 illustrates drop-out rates for 7th- to 12th-graders in the comparison group of Pennsylvania school districts. Philadelphia and Harrisburg have the highest drop-out rates among schools in the Pennsylvania comparison group, and Cheltenham and Abington have the lowest rates.

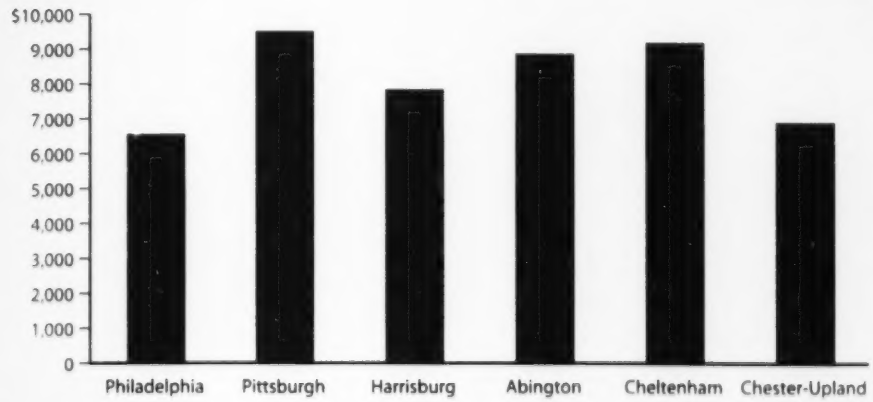
Teacher Absences

In addition to Philadelphia and Harrisburg having the highest drop-out rates of any school district in the Pennsylvania comparison group, teachers in these two school districts are more likely than teachers in other schools in the Pennsylvania comparison group to be absent for personal reasons. As figure 10 shows, on any given day 6.5 percent of Philadelphia teachers and 6.8 percent of Harrisburg teachers were absent for personal reasons, compared with 5 percent of Pittsburgh teachers, 4.7 percent of Cheltenham teachers, and 4.9 percent of Abington teachers. (Chester-Upland did not report the percentage of teachers absent for personal reasons.) Although there is no reason to believe that these absences are not legitimate, the high absence rate in Philadelphia may indicate that Philadelphia teachers are somewhat less committed than teachers in Pittsburgh or the suburban systems to arranging their personal lives so that they are at school as much as possible. The high absence rate of Philadelphia teachers is certainly a topic worthy of further investigation.

Student Achievement

Not surprisingly, students in the two middle class suburban districts in the sample, Cheltenham and Abington, scored better, on average, than students in Philadelphia on statewide assessments in mathematics, reading, and writing given in 1997.[2] Although Philadelphia Public School students had average test scores in 1997 on statewide assessments that were lower than the statewide average, their average scores are comparable to those of students in Harrisburg and higher than those of students in the Chester-Upland School District (figure 11). However, students in Pittsburgh scored, on average, almost 100 points higher on the 1997 statewide assessment than students in Philadelphia and Harrisburg, despite the fact that Pittsburgh's percentage of low-income children was comparable to Harrisburg's and only slightly lower than Philadelphia's that year (figure 6). As discussed above, Pittsburgh's drop-out rates were also lower than Philadelphia's or Harrisburg's in 1996–97. Although it is

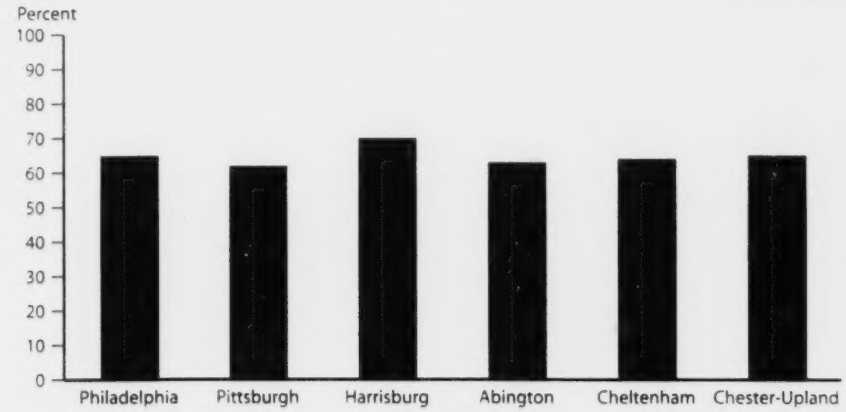
Figure 7.—Current expenditures per pupil for selected Pennsylvania school districts, 1995–96



SOURCE: Pennsylvania System of School Assessment School Profiles 1996–97.

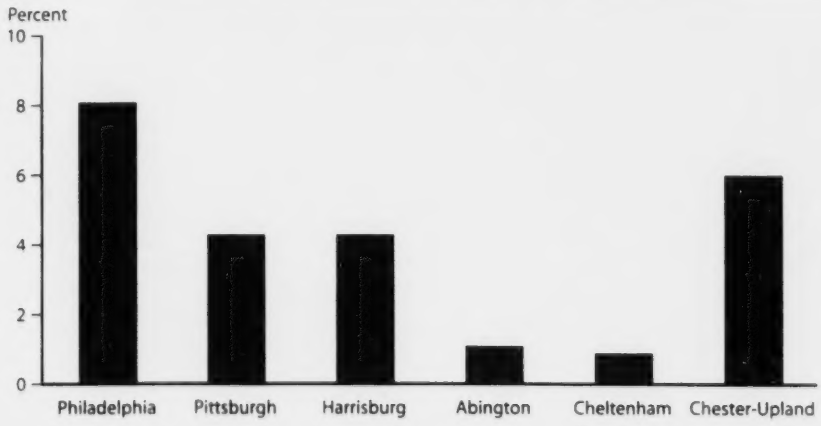
[2] We were not able to obtain standard errors for the Pennsylvania state tests, so we cannot determine if differences are statistically significant.

Figure 8.—Percentage of expenditures going toward instructional activities for Pennsylvania school districts, 1995–96



SOURCE: Pennsylvania System of School Assessment School Profiles, 1996–97.

Figure 9.—Drop-out rates for selected Pennsylvania school districts, 1996–97



SOURCE: Pennsylvania System of School Assessment School Profiles, 1997–98.

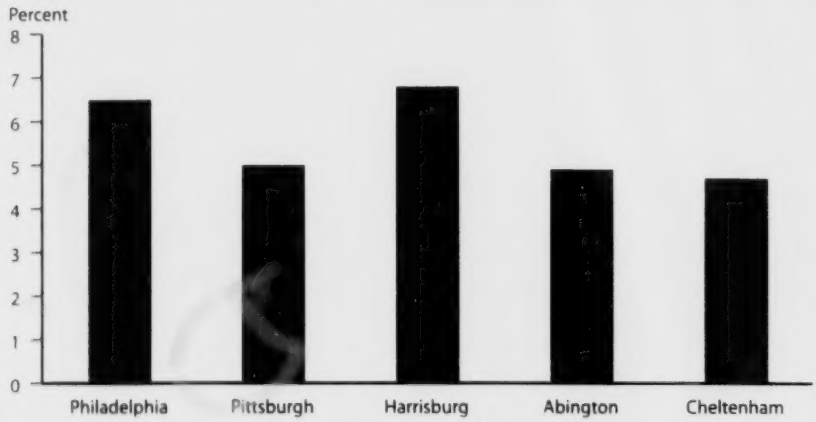
extremely difficult, if not impossible, to draw causal relationships between spending per pupil and educational outcomes, and the data we have presented here certainly do not allow us to do so, the fact that Pittsburgh spent $3,000 more per pupil than Philadelphia and $1,300 more per pupil than Harrisburg in 1996–97 indicates that the relationship between spending and educational outcomes should be further investigated in this case (figure 7).

Comparisons between Philadelphia and the State of Pennsylvania, and Other Big City School Districts and States

Comparisons between Cities and States

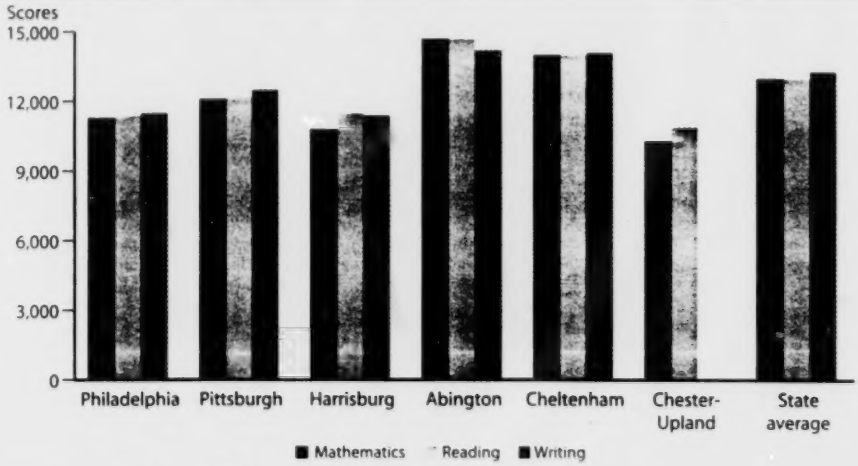
Although each state measures such indicators as per pupil spending and drop-out rates somewhat differently, these measurements are usually consistent for all school districts within a state. Thus, if one school district in a

Figure 10.—Percentage of teachers absent due to personal reasons for selected Pennsylvania school districts, 1995–96



SOURCE: Pennsylvania System of School Assessment School Profiles, 1996–97.

Figure 11.—Grade 11 mathematics and reading and grade 9 writing assessment scores for selected Pennsylvania school districts, 1995–96



NOTE: No writing assessment data is available for Chester-Upland School District.

SOURCE: Pennsylvania System of School Assessment School Profiles, 1996–97.

state spends twice as much as another, we can be reasonably confident that each school district is measuring the same set of expenditures, and the higher spending district does spend twice as much as the lower spending district. If the school districts are in different states, we cannot assume that they are measuring the same set of expenditures. It is possible that the school district spending twice as much is including expenditures in its total, such as capital expenditures, that the lower spending school district in a different state does not include.

By reporting a ratio of city spending to state spending, we can compare cities in different states, despite the fact that states measure spending differently. For example, if one city spends more than its state average, and another city spends less than its state average, we can infer that the first city is probably making a bigger effort to meet the needs of its students than the second city.

Similarly, states measure drop-out rates differently, but cities within a given state usually measure drop-outs rates

consistently. Some states report drop-out rates for 7th- to 12th-graders, some states report drop-out rates for 9th- to 12th-graders, and other states report drop-out rates for 1st- to 12th-graders. The drop-out rate for 1st- to 12th-graders is always lower in any given school district than the drop-out rate for 9th- to 12th-graders, because so few students drop-out of school before 9th grade.

By reporting a ratio of city drop-out rates to state drop-out rates, we can compare cities in different states. If one city has a drop-out rate twice its state average, and another city has a drop-out rate equal to its state average, the first city probably has a bigger drop-out problem, even if the two states measure drop-out rates somewhat differently.

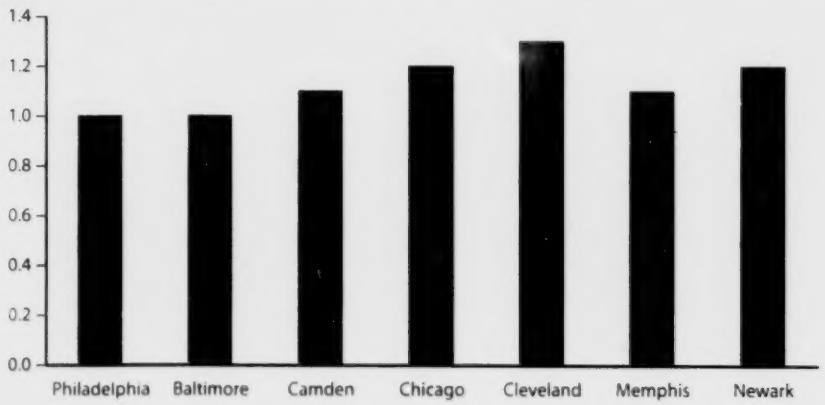
Spending Per Pupil

Philadelphia is one of only two big city school districts in our comparison group that does not spend more per pupil than the statewide average (figure 12). Generally, expenses are higher for big city school districts than for the state as a whole, because cities usually have higher labor costs than suburban or rural school districts. In addition, the pupils in city schools usually have higher needs than the pupils in rural or suburban schools. As illustrated in figure 2, most city schools have high percentages of children living below the poverty line whose educational needs are often great. Therefore, it is troubling that Philadelphia is spending only the state average per pupil.

Ratio of City School District Administrative Expenditures to Instructional Expenditures

As illustrated in figure 13, Philadelphia's ratio of administrative to instructional expenses is low among the comparison cities for which we were able to obtain data. Philadelphia's spending on administrative expenses is equal to approximately 11 percent of its spending on instructional expenses. Among our comparison cities, only Baltimore has a lower ratio of administrative to instructional expenditures.

Figure 12.—Ratio of average per pupil spending in city school districts to average per pupil spending in states



NOTE: Ratios greater than 1.0 indicate average per pupil spending by a city school district is greater than average per pupil spending at the state level. Ratios equal to 1.0 indicate that average per pupil spending by a city equals average per pupil spending at the state level. Philadelphia: Figures may not be exact due to incomplete reporting. Baltimore: Cost per pupil reflects the average cost of providing educational and related services to the students of the local school system. Philadelphia, Baltimore, and Chicago: Data reflect the 1996–97 school year. Camden: Data reflect the 1997–98 school year. Cleveland: Data reflect FY97 figures. Memphis: Data reflect 1996–97 Operating Expenditures per Student. Newark: Data reflect the 1997–98 school year.

SOURCE: Philadelphia: Data taken from 1996–97 Pennsylvania Department of Education database. Baltimore: Data taken from 1997–98 Maryland State Department of Education Fact Book. Camden: Data taken from the New Jersey Department of Education Comparative Spending Guide. Chicago: Data taken from the 1997 School Report Card issued by the Illinois Board of Education. Cleveland: Data taken from the Cleveland City School District Profile distributed by the Ohio Department of Education. Memphis: Data taken from the Tennessee Department of Education 1997 Report Card. Newark: Data taken from the New Jersey Department of Education Comparative Spending Guide.

Teacher/Pupil Ratio

In addition to spending as much or more per pupil as the state average, the city school districts in our comparison group have teacher:pupil ratios that are as high, or slightly higher, than the state average (figure 14). Philadelphia's teacher/pupil ratio is approximately 10 percent higher than the state of Pennsylvania's teacher:pupil ratio. Both Cleveland and Memphis have higher teacher:pupil ratios in relation to their state average than Philadelphia. Cleveland has a teacher:pupil ratio that is 20 percent higher than the Ohio average, and Memphis has a teacher:pupil ratio that is 30 percent higher than the Tennessee average (figure 14).

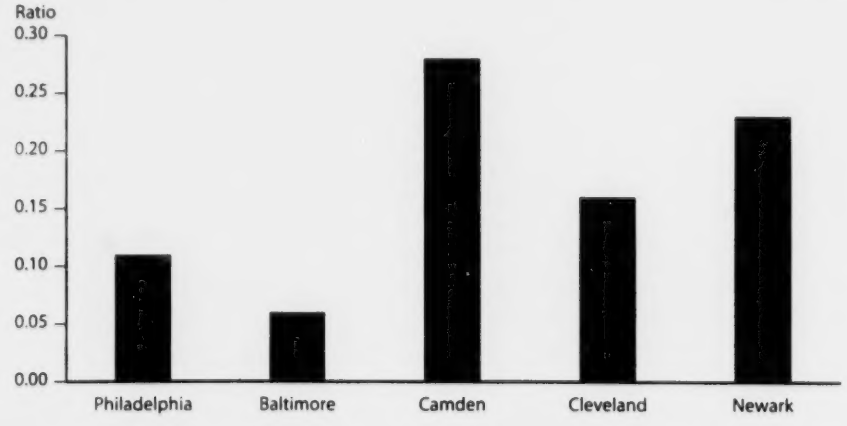
Drop-out Rates

Approximately three students in Philadelphia drop out of school before finishing 12th grade for each student in the state of Pennsylvania who drops out of school before finishing 12th grade (figure 15). Among the schools in our comparison group for which we were able to obtain drop-out rate data, only Cleveland has a higher ratio of city drop-out rate to state drop-out rate.

Conclusion

None of the factors used to compare Philadelphia with other large, high poverty cities indicates that the Philadelphia Public School District is doing a significantly worse job than the school districts in these other cities. Philadelphia's student/teacher ratio is at the high end for this group, but it is not the highest. Philadelphia's current expenditures per pupil are in the middle of the group of larger city comparison districts. The fact that Philadelphia's average SAT scores are in line with those in other large cities in the comparison group indicates

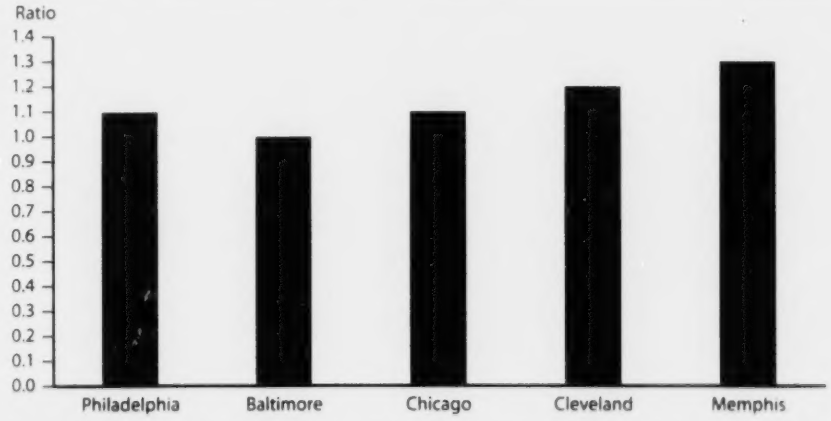
Figure 13.—Ratio of city school district administrative expenditures to instructional expenditures



NOTE: Data for Chicago Public Schools and Memphis City School District are not available. Philadelphia: Data from the 1996–97 school year. Baltimore: Administration includes system wide regulation, direction, and control of the Local Education Agency (LEA), including office of the superintendent, business services, centralized support services, and instructional direction and improvement services. Instruction includes activities that address teaching regular students or enhancing the educational experience for students. Included in this category are classroom instruction, excluding special education services; school media services; cocurricular activities; office of the principal; guidance services; and psychological services. Camden: Data reflect the 1997–98 school year. Cleveland: Data reflect FY97 figures. Newark: Data reflect the 1997–98 school year.

SOURCE: Philadelphia: Data taken from Pennsylvania Department of Education for the 1996–97 school year. Baltimore: Data taken from Maryland State Department of Education for the 1996–97 school year. Camden: Data taken from New Jersey Department of Education Comparative Spending Guide. Cleveland: Data taken from the Cleveland City School District Profile distributed by the Ohio Department of Education. Newark: Data taken from New Jersey Department of Education Comparative Spending Guide.

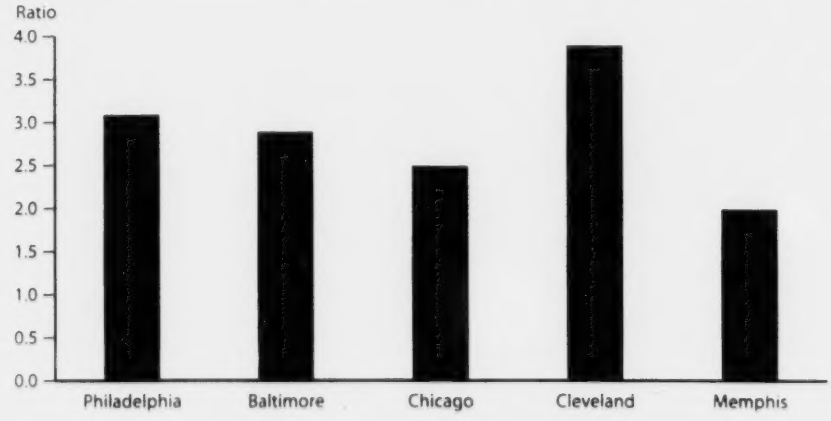
Figure 14.—Ratio of city school district teacher/pupil ratio to state teacher/pupil ratio



NOTE: Data for Camden City School District and Newark City School District are not available. Philadelphia: Data include teachers at all levels and in all areas. Cleveland: Data reflect FY97 figures.

SOURCE: Philadelphia: Data taken from 1996–97 figures regarding total number of teachers and total enrollment. This includes teachers at all levels and in all areas. Baltimore: Data taken from 1996–97 full-time-equivalent enrollment figures provided by the Maryland Department of Education. Chicago: Data taken from the 1997 School Report Card issued by the Illinois Board of Education. Cleveland: Data taken from the Cleveland City School District Profile distributed by the Ohio Department of Education. Memphis: Data taken from the Tennessee Department of Education 1995–96 School Year Annual Statistical Report.

Figure 15.—Ratio of city school district drop out rates to state drop out rates



NOTE: Data for Newark City Schools and Camden City Schools are not available. Philadelphia: Figures reflect the annual rate. Baltimore: Drop-out rate reflects the rate for students grades 9–12 grade for the 1996–97 school year. Cleveland: Data reflect FY97 figures. Memphis: Data reflect 1996–97 dropout rates for students in public schools, grades 9–12.

SOURCE: Philadelphia: Data taken from the Pennsylvania Department of Education 1996–97 data base. Baltimore: Data taken from 1997–98 Maryland State Department of Education Fact Book. Chicago: Data taken from the 1997 School Report Card issued by the Illinois Board of Education. Cleveland: Data taken from the Cleveland City School District Profile distributed by the Ohio Department of Education. Memphis: Data taken from the Tennessee Department of Education 1997 Report Card.

that the students in Philadelphia who are considering applying to college are receiving an education that is comparable with those of students in the other large cities as measured by the SAT examination.

On average, large city schools do have students with greater needs than students in other schools in their state. Socioeconomic status has been repeatedly shown to be highly correlated with student achievement, and large city schools generally have students with lower average socioeconomic status than suburban schools. Thus, the fact that Philadelphia is one of only two big city school systems in the comparison group that spends less than the state average per pupil is somewhat troubling, as it indicates that all the needs of Philadelphia's students may not be adequately met. However, the fact that Philadelphia is spending a lower percentage of its budget on administration than all but one of the other big city schools indicates that expenditures that directly impact students in the classroom may not be as deficient as they first appear.

As expected, Philadelphia appears in a less favorable light when compared with wealthier suburban Pennsylvania school districts. Philadelphia's drop-out rate is higher, and its test scores are lower, than the two neighboring districts in the comparison group. Philadelphia also has a higher teacher absentee rate than these districts, which can be an indicator of teachers' lack of commitment to their job and students.

However, when compared with a high poverty neighboring suburban district, Philadelphia did not appear to perform nearly so poorly. While Philadelphia's drop-out rate was higher than in the neighboring high poverty suburban school, its average test scores on the Pennsylvania state tests were also higher.

Finally, when compared with the other two large cities in Pennsylvania, Pittsburgh and Harrisburg, Philadelphia's performance did not appear unusually poor. Students in Philadelphia have average test scores comparable with Harrisburg, although they were lower than the average test scores in Pittsburgh. Philadelphia's drop-out rates and percentage of teachers absent for personal reasons were also comparable with Harrisburg's and lower than Pittsburgh's.

While the data discussed above indicate that there is probably substantial room for improvement in the Philadelphia public schools, both in terms of inputs such as expenditures per pupil, and in terms of outcomes such as drop-out rates and test scores, the data do not indicate that Philadelphia's performance is substantially divergent from what one would expect to see in a large, high poverty school district. Drastic measure may indeed be required to improve the performance of the Philadelphia public schools, but if they are, schools in other cities also require similar intervention.

School Choice in Milwaukee: Are Private Schools Creaming Off the Best Students?

Dan D. Goldhaber
The Urban Institute

Dominic J. Brewer
RAND

Eric R. Eide
Brigham Young University

About the Authors

Dan D. Goldhaber is a labor economist who serves as a Research Associate at the Urban Institute in the Education Policy Center and as a member of the Alexandria City School Board. His research focuses on educational productivity and reform at the K–12 level and on teacher labor markets. Examples of work in these areas include analyses of the demographic and productivity impacts of educational vouchers, the effects of teacher qualifications and quality on student achievement, and the effect of the interaction between student and teacher race, gender, and ethnicity on student outcomes. He has published in numerous academic economics and education journals including the *Journal of Human Resources*, the *Journal of Urban Economics*, *Economics of Education Review*, *Education Economics*, *Industrial and Labor Relations Review*, and *Phi Delta Kappan*. Dr. Goldhaber has served as a reviewer for numerous academic journals and presented his research at professional meetings such as the American Economic Association, the American Educational Research Association, the Association for Public Policy and Management, and the American Educational Finance Association. Dr. Goldhaber received a Ph.D. and M.S. in Labor Economics from Cornell University and a B.A. in economics from the University of Vermont.

Dominic J. Brewer is a labor economist at RAND specializing in the economics of education and the Director of RAND Education. In recent years his research has focused on educational productivity and teacher incentives in both K–12 and higher education. His research has examined policy issues through the analysis of large national databases including the original *Coleman Report* data, *High School and Beyond* and the *National Educational Longitudinal Study of 1988*. Examples of this work include an analysis of the effects of teacher education and quality on student achievement gains, the interaction between student and teacher race, gender and ethnicity, and the effects of administrative resources on student performance. He has completed a series of studies on the effects of ability grouping on student achievement using an American Education Finance Association/National Center for Education Statistics Young Scholars Award. Recent higher education research has included a study of community college faculty's connections to the labor

BEST COPY AVAILABLE

market for the National Center for Research in Vocational Education, and a study of the labor market payoff to attending different types of four-year college. He has published in numerous academic economics and education journals including *Review of Economics and Statistics, Educational Evaluation and Policy Analysis, Journal of Labor Economics, Journal of Human Resources,* and *Journal of Policy Analysis and Management.* He is an associate editor of *Economics of Education.* He serves on the National Center for Education Statistics Research Design Panel on school finance. Dr. Brewer received a Ph.D. in Labor Economics from Cornell in 1994, and holds a bachelor's degree from Oxford. He has been at RAND since 1994 and has also been a Visiting Assistant Professor of Economics at UCLA and on the faculty of the RAND Graduate School.

Eric R. Eide is an associate professor in the economics department at Brigham Young University. He is a labor economist whose research focuses on the economics of education and earnings inequality. His early research in higher education examined how college major affects an individual's lifetime earnings, as well as how differences in college major choice contribute to earnings differentials among various gender and racial groups. Professor Eide has more recently analyzed higher education issues such as the labor market premium associated with attending colleges of differing selectivity, and the effect of college education on earnings inequality in the United States. Examples of his work on primary and secondary education include studies on the long run consequences of grade retention, the effect of school spending on the distribution of student achievement and labor market earnings, and the relationship between participating in extracurricular activities and educational attainment and labor market outcomes. His research has been published in economics journals such as the *Journal of Human Resources, Economics of Education Review, Journal of Population Economics, Southern Economic Journal, Economics Letters,* and *Contemporary Economic Policy.* Professor Eide has been on the faculty at Brigham Young University since 1993, where he has taught courses on labor economics, econometrics, and statistics. He received a Ph.D. in Economics from the University of California, Santa Barbara in 1993, and he completed bachelor's and master's degrees at Brigham Young University.

School Choice in Milwaukee: Are Private Schools Creaming Off the Best Students?

Dan D. Goldhaber
The Urban Institute

Dominic J. Brewer
RAND

Eric R. Eide
Brigham Young University

Introduction

There has been considerable debate over the results of the voucher program that has been in place in Milwaukee since 1990. Researchers who have analyzed the effect of participating in the Milwaukee Parental Choice Program (MPCP) on student mathematics and reading test scores have reached different conclusions. One possible explanation for the divergent findings is that there are important differences in student characteristics between the public and private sectors; namely, that private schools are creaming off students from the public schools who are most likely to have high academic achievement, and that researchers examining this issue have used different methodologies to account for this. Some of the differences in student characteristics, such as family income and parental education, are readily observable and can easily be incorporated into standard statistical models. However, in the case of private schooling, statisticians worry about "sample selection"—the degree to which "unobservable" differences between students in the two sectors generate biased estimates of the effects of private schooling.

Sample selection occurs when characteristics of students that influence academic achievement and are unobservable in the data are systematically related to the school sector in which they are enrolled.[1] We might expect sample selection to be present in the case of private schooling because parents freely choose the school sector in which their children are enrolled. This fact suggests that parents who choose the private sector may be quite different from other parents, but in ways that are difficult to identify. For instance, in most cases parents have to pay additional monies, both in tuition and in transportation costs, to send their children to private schools. These parents demonstrate a willingness to support education which could indicate that they also provide an environment in the home which is conducive to educational achievement; factors which are difficult to account for statistically.

If there are important "unobservable" differences, then standard statistical models (ordinary least squares) of achievement are inadequate. The problem faced is that

[1] See Goldhaber (1996) and Figlio and Stone (1997) for a detailed discussion of sample selection associated with private schooling.

BEST COPY AVAILABLE

it cannot be determined whether differences in achievement between public and private school students are due to genuine differences in performance attributable to school type or to underlying differences in motivation, home environment, etc. Thus, central to the debate over the effects of school choice is the issue of sample selection.

Data from the Milwaukee school choice experiment allow researchers to explore the issue of selection because there are public school students who had the opportunity to apply to participate in a voucher program but chose not to, and there are students in the public sector who applied for the voucher program but were rejected from it. Here, we examine the Milwaukee data to investigate the extent to which those students who enrolled in the Milwaukee choice program differed from their public school counterparts in terms of both "observed" and "unobserved" characteristics. We begin with a discussion of the school choice program in Milwaukee.

School Choice In Milwaukee

The Milwaukee Parental Choice Program began in the fall of 1990. The parameters of the MPCP are as follows. Students enrolled in the Milwaukee public school (MPS) district who came from families with incomes not exceeding 1.75 times the national poverty line, and who were not enrolled in a private school in the immediate prior year, were eligible to attend private, nonsectarian schools in the district (Witte 1997). The total number of choice students in any year was limited to 1 percent of the MPS membership in the first four years, and was increased to 1.5 percent for the 1994–95 school year.[2] For each choice student who enrolled, their respective private school received a payment equivalent to the MPS per student state aid (about $3,200 in 1994–95). Private schools were required to limit choice students to 49 percent of their enrollment (this figure rose to 65 percent beginning in the 1994–95 school year) and schools that were oversubscribed were required to accept students based on a random selection.[3] This last provision provides a "natural experiment" because several of the schools were oversubscribed, resulting in the random assignment of students between the public and private sectors. In other words, in several cases there was a greater demand for participation in private schools in the MPCP than there were slots available (given the specifications of the program) in those schools.[4] The randomness of program participation allows researchers to compare students who applied for admission through the MPCP, were not selected through the lottery process, and therefore attended a public school, to those who applied through the MPCP, were selected through the lottery process, and attended private school.[5] In theory, this approach avoids the selection problem discussed in the previous section.

[D]ifferences in achievement between public and private school students are due to genuine differences in performance ... or to underlying differences in motivation, home environment, etc.

A great deal of effort has been spent collecting student-level data for the evaluation of the Milwaukee experiment. Family background information was solicited from all participating students and for a large random sample of nonparticipating public school students, and tests were administered in the spring to both program participants and nonparticipants in grades K–8.[6] The tests administered are the Iowa Tests of Basic Skills (ITBS) which is a nationally normed standardized test with scores ranging from 1 to 99 with a national mean of 50. Several researchers have examined these data to determine the relative effectiveness of public and private schools in Milwaukee.

[2] The program began with an enrollment of 341 students in 7 schools. By 1995, enrollment had risen to 830 students, with 12 schools participating in the program (Witte 1997).

[3] Schools were required to admit choice students without discrimination based on race, ethnicity, or prior academic performance, but were not required to admit disabled students (Witte 1997).

[4] The school choice program was amended in June 1995. For details of the changes to the program see Witte (1997).

[5] However, as Witte (1997) notes, there was no authority ensuring that the selection was, in fact, random.

[6] In total, approximately 4,000 student-year observations were collected. However, students who were observed in multiple years contributed more than a single observation.

BEST COPY AVAILABLE

John Witte (1997), who was commissioned by the state of Wisconsin to evaluate the Milwaukee program, compares the students who enrolled in the MPCP with a sample of students enrolled in the MPS system. Although he finds that the parents of students enrolled in a private school were generally more satisfied with their schools, there was no case in which private schools outperformed public schools, controlling for the family background of students.

Greene et al. (1998) use an alternate methodology, comparing students who participated in the MPCP with those who applied through the MPCP to attend a private school but were rejected due to over subscription and thus attended a MPS. In theory, the comparison of "selected" private school students and those students who applied for the program but were nonselected based on a lottery (the "lottery not selected"), avoids the problems associated with sample selection. The hypothesis is that the "lottery not selected" do not have unobservable characteristics that are systematically different from those who applied to the program and were accepted. Greene et al. (1998) find little evidence of a private school effect for students in the first two years of the program but a large private school advantage in years three and four. For instance, controlling for family background, Greene et al. (1998) estimate a statistically significant private school advantage on standardized tests of 7 percentile points in mathematics in year three, and 6 percentile points in reading year three of the program.[5]

Rouse (1997) uses instrumental variable estimation techniques to address the possibility that selection bias exists due to the fact that not all those who are selected to participate in the MPCP actually attend, and program participation is not likely to be random. The instrument is the initial selection into the program, which is assumed to be correlated with attendance at a choice school, but uncorrelated with achievement. The analysis finds that students who attended a choice school scored between 1 and 2 percentage points per year higher in mathematics than students who were not selected. The methodology used by Greene et al. (1998) and Rouse (1997) in which choice participants and "lottery not selected" are compared, does not allow for inference outside of those students who applied to participate in the MPCP.[6] In other words, the effect of attending a private school may be different for the average student and the student who chooses to apply to the MPCP.

Witte and Thorn (1996) examine the type of students who participate in the MPCP. They find that choice parents were more likely to be involved with their children's schooling prior to participating in the MPCP, rated their prior school lower, and have higher educational expectations for their children than do nonparticipants. These results suggest that choice participants may differ from nonparticipants in important ways that are difficult to empirically quantify.

Methodology and Data

We can empirically test what observable characteristics affect the probability of applying to the MPCP by estimating a probit regression of the binary decision to apply to the program:

$$D_i = \gamma V_i + u_i$$
and define $D_i = 1$ *if* $D^*_i > 0$ (applied for voucher) (1)
$D_i = 0$ *if* $D^*_i \leq 0$ (did not apply)

V is a vector of variables assumed to affect the choice to enroll in the program including student race and gender, initial student test scores (tests taken in the fall prior to the decision to apply) in reading and mathematics, student grade level, family income and parental education, whether the student is from a single parent family, whether the family is Catholic, whether a foreign language is spoken in the home, and the distance to the nearest choice school. γ is the effect of these variables on the probability of applying to the MPCP.

Students who are rejected from the MPCP may elect to attend an alternate private school, in which case they do not appear in the sample, or attend the Milwaukee public schools. We use these students to test for the existence of sample selection associated with the decision to participate in the choice experiment. To determine whether unobservable differences exist between applicants and nonapplicants, we estimate achievement models for

[5] Both the year three mathematics and reading results are only statistically significant at the 10 percent level for one-tail tests.

[6] This point is also made by Rouse.

BEST COPY AVAILABLE

public school students who applied for the MPCP (and who were rejected from the program) and nonapplicants, which allow the coefficients between the two groups of students to vary. Because both groups of students remained in the public school system, in theory, the coefficients of the two groups should not be systematically different.

More formally, letting A_{iK} represent the achievement of student i in group K (the applicant or nonapplicant group) and X_i represent the individual and family background characteristics for student i, we can estimate the following achievement model:

$$A_{iK} = \beta X_i + B_K (X_i \ D_i) + \varepsilon_i \qquad (2)$$

We test whether unobservables affect the returns to observable characteristics by examining the B_K coefficients to determine whether the differences in returns between the two groups are jointly equal to zero—that is, whether H_o: $B_K = 0$. This is a straightforward F-test to see if a restricted model that does not allow the coefficients to differ between groups is significantly different from the more flexible specification. Rejection of the null hypothesis indicates that unobservables systematically differ between applicants and nonapplicants.

The sample of students that we examine is restricted such that it corresponds to those students who were eligible to participate in the MPCP. Most students in the MPS were tested only in the second, fifth, and seventh grades. Only those public school students who qualified for Title I aid were required to be tested every year (Title I status roughly corresponds to eligibility for the voucher given; in order to qualify for free and reduced-price lunch a family had to be below 185 percent of the poverty line). We restrict our sample to those who were tested every year, thus most MPS students who were not eligible for the MPCP are excluded from our sample.[9]

Table 1 lists selected sample statistics for the dataset that we analyze. The table is divided into three categories: nonapplicant public school students, public school students who applied for the MPCP but were rejected, and applicants who were accepted into a private school.

Public school nonapplicants score better in both the mathematics and reading achievement tests, come from families with higher family income, and have more educated parents than public school applicants and private school students. Private school students have larger year-to-year gains in mathematics (judging from the difference between the pre- and post-mathematics test) than do public school applicants and nonapplicants. By contrast, private school reading scores actually fall from year-to-year.

It is interesting to note that applicants to the MPCP are more likely to be from single parent families. Although public school students (both those who applied and those who did not apply to the MPCP) gained more than a point on the reading test score, private school choice students scored lower in the spring on the reading test than they did in the fall.

Results

To investigate which observable characteristics affect the probability of applying to the MPCP, we estimate a binary probit model. The coefficient estimates from this model are shown in table 2.

We find that students from lower income families were significantly more likely to apply to the MPCP, as were black and Hispanic students. Students in higher grades are less likely to apply to the program (as evidenced by the negative coefficient on current grade level), and interestingly, the coefficient on Catholic was insignificant indicating that Catholics were no more or less likely to apply. Finally, the coefficient on the distance variable is negative and statistically significant suggesting that, as expected, the further a student lives from a choice school, the less likely it is the student will apply for the MPCP.

In some respects the MPCP allowed private schools to attract a "select" group of students from the public schools. For instance, we find students with more educated parents (those who had at least one parent who went to college or at least one parent who went to high school) were more likely to apply. However, by many measures, no creaming is occurring. Students from lower incomes were more likely to apply to the program as were those who

[9] One caveat is that schools with a large majority of Title I students sometimes tested the whole school. Thus, it is possible that some students not eligible for the MPCP who were attending high poverty schools, which elected to test the entire student population, are included in this sample.

Table 1.—Sample means

Variables	Nonapplicants public school	Applicants: Nonselected public school	Applicants: Choice program private school
Base year reading score	42.44 (17.53)	36.16 (13.88)	39.60 (16.50)
Post reading score	43.45 (17.23)	36.78 (15.99)	38.71 (14.89)
Base year mathematics score	44.56 (20.50)	37.36 (19.83)	41.29 (18.68)
Post mathematics score	44.73 (20.45)	38.91 (21.91)	42.31 (17.25)
Female	0.506	0.524	0.540
Black	0.622	0.837	0.821
Hispanic	0.127	0.115	0.129
Catholic	0.217	0.180	0.125
Current grade level	3.72 (1.94)	3.23 (1.86)	3.40 (1.96)
Family income	14,586 (12,361)	9,823 (7,559)	11,609 (7,689)
Single parent household	0.639	0.754	0.747
Parent graduated college	0.080	0.033	0.092
Parent graduated high school	0.750	0.787	0.846
Foreign language spoken at home	0.022	0.033	0.012
Sample size	1,392	61	487

NOTE: Standard deviations in parentheses.

SOURCE: Author's calculations.

performed poorly in mathematics. This suggests that, contrary to some private schools, the MPCP schools are not enrolling mainly the elite students from the MPS system.

The results of the probit regression clearly show that several observable characteristics are correlated with the decision to apply to the MPCP. To test whether there are unobservable characteristics that are correlated with the decision to apply to the MPCP, we compare public school students who applied for the MPCP and were rejected to public school students who were eligible for the program but did not apply. We estimate models in which the achievement on mathematics and reading tests is modeled as a function of student and family background variables, including the previous year's achievement in mathematics and reading.[10] In these equations, we allow the returns to observable characteristics to vary by application status by allowing the intercept for each category of students to differ, and by including interaction terms between application and the observable characteristics that allow the coefficients to differ. Since all students in these models (those who applied to the program and were rejected and those who did not apply) went to Milwaukee public schools, in theory, there should be no significant differences in the returns to individual and family background characteristics. The existence of differences in these estimated coefficients provides evidence that there are unobservable variables which are correlated with the decision to apply to the program, and thus, sample selection exists.

Table 3 shows the estimated coefficients of the yearly mathematics and reading achievement models. Columns (1) and (3) are the estimates, for mathematics and reading, respectively, when the intercept is allowed to differ. In both mathematics and reading, parental education and base year test scores, both within and across subjects, positively affect achievement. Also, it is interesting to note that the negative statistically significant coefficient, in both subjects (on current grade), implies that as students progress through the MPS system, they fall behind relative to the rest of the country. The intercept was not found to differ between applicants and nonapplicants

[10] Our sample includes multiple observations of individual students. We estimated additional models, in which the sample was restricted to one observation of each student, and the results were not substantively different. We report heteroskedasticity-corrected standard errors.

Table 2.—Probit of application to the choice program		
Intercept	-1.055*	(0.160)
Distance (in miles) to the nearest choice school	-0.025*	(0.011)
Base year mathematics score	-0.003*	(0.001)
Base year reading score	-0.002	(0.002)
Foreign language spoken in home	-0.271	(0.202)
Family income (in thousands)	-0.014*	(0.003)
Catholic	-0.094	(0.089)
Single parent household	0.029	(0.068)
Female	0.083	(0.053)
Black	0.882*	(0.095)
Hispanic	0.964*	(0.120)
Current grade level	-0.067*	(0.013)
Parent graduated college	0.393*	(0.107)
Parent graduated high school	0.371*	(0.071)
Log likelihood		-1,524
Sample size		3,050

* Variable is significant at the .05 level.

NOTE: Standard errors in parentheses.

SOURCE: Author's calculations.

which provides cursory evidence that there are no significant unobservable effects that are correlated with application status.

Columns (2) and (4) of table 3 show the results, for mathematics and reading, when the more flexible specifications of the model are estimated. Several of the interaction terms are significantly different from zero. In mathematics, those applicants to the MPCP had a higher return to their initial level of mathematics achievement (the base mathematics score), and single-parent households who are choice applicants do better on the mathematics achievement test than do nonapplicants. Likewise, applicant children from single-parent households do better in reading than do nonapplicants (although this variable is only significant at the 0.10 level).

Despite these differences in individual coefficients, on the mathematics and reading tests, we could not reject an F-test (at the 95 percent confidence level) of the hypothesis that the interaction terms are jointly equal to zero. However, although it does not meet the generally accepted statistical significance level of 95 percent, in mathematics we could reject the null hypothesis at about the 90 percent confidence level. This provides some indication that applicants differ systematically from nonapplicants in terms of their unobservable characteristics.

Conclusion

In this paper, we have utilized the unique nature of the data from the MPCP experiment to shed some light on the question of whether students applying for the program have characteristics that systematically differ from students who do not apply. We find students applying to the MPCP differ markedly from nonapplicants in terms of observable characteristics. However, in reading, there is no evidence suggesting systematic differences in unobservables between the two groups of students, and, in mathematics, there is only weak evidence that differences exist. Hence, overall, our findings indicate there is little evidence of sample selection associated with the decision to participate in the choice experiment.

These results indicate the divergent findings of earlier research on the achievement effects of enrollment in the Milwaukee choice program cannot readily be explained by the research methods used to control for unobservables between applicants who attend private schools in Milwaukee and nonapplicant MPS students.

Table 3.—Ordinary Least Squares (OLS) estimates of achievement for public school students

	Mathematics		Reading	
	(1)	(2)	(3)	(4)
Intercept	19.110*	19.548*	16.634*	16.958*
	(1.992)	(2.008)	(1.771)	(1.793)
Base year mathematics score	0.584*	0.575*	0.168*	0.167*
	(0.024)	(0.024)	(0.022)	(0.022)
Base year reading score	0.117*	0.125*	0.483*	0.484*
	(0.028)	(0.029)	(0.025)	(0.025)
Foreign language spoken in home	-3.066	-2.769	-0.148	-0.304
	(2.653)	(2.720)	(2.295)	(2.355)
Family income (in thousands)	0.073	0.065	0.036	0.036
	(0.039)	(0.039)	(0.035)	(0.034)
Catholic	0.416	0.083	0.401	0.393
	(1.100)	(1.127)	(0.993)	(1.012)
Single-parent household	0.275	-0.147	0.347	0.073
	(0.922)	(0.933)	(0.817)	(0.830)
Female	0.661	0.785	1.898*	1.753*
	(0.749)	(0.760)	(0.668)	(0.679)
Black	-4.885*	-4.976*	-4.173*	-4.078*
	(1.074)	(1.083)	(0.963)	(0.974)
Hispanic	-2.585	-2.663	-2.991*	-3.052*
	(1.414)	(1.423)	(1.265)	(1.277)
Current grade level	-1.336*	-1.305*	-0.454*	-0.450*
	(0.196)	(0.198)	(0.174)	(0.177)
At least one parent went to college	3.080*	3.022*	3.277*	3.031*
	(1.594)	(1.612)	(1.405)	(1.423)
At least one parent went to high school	1.479	1.453	1.406	1.268
	(0.928)	(0.941)	(0.825)	(0.840)

Table 3.—Ordinary Least Squares (OLS) estimates of achievement for public school students–Continued

	Mathematics		Reading	
	(1)	(2)	(3)	(4)
Interactions with application for the Milwaukee Parental Choice Program				
Intercept	-0.518 (2.076)	-27.289 (15.054)	-1.631 (1.894)	-10.050 (13.092)
Base year mathematics score		0.450* (0.175)		0.074 (0.154)
Base year reading score		-0.386 (0.205)		-0.117 (0.189)
Foreign language spoken in home		-12.318 (13.898)		-4.008 (12.986)
Family income (in thousands)		0.447 (0.335)		0.060 (0.291)
Catholic		5.592 (7.405)		-1.692 (8.590)
Single-parent household		16.828* (6.571)		10.616 (5.872)
Female		-7.635 (5.179)		2.080 (4.532)
Black		2.358 (11.982)		-9.953 (10.770)
Hispanic		12.782 (15.438)		-2.343 (14.301)
Current grade level		1.217 (1.529)		1.052 (1.195)
Parent graduated college		10.022 (13.629)		13.007 (12.792)
Parent graduated high school		3.656 (5.97[illegible]		6.725 (5.467)
F-statistic of the joint significance of the interaction terms		[illegible]69		0.676
Prob > F		0.114		0.788
Adjusted R^2	0.544	[illegible]46	0.489	0.488
Sample size	1,431	1,431	1,440	1,440

* Variable is significant at the .05 level.

NOTE: Standard errors in parentheses.

SOURCE: Author s calculations.

References

Figlio, D. N., and Stone, J. A. 1997. "School Choice and Student Performance: Are Private Schools Really Better?" Working paper.

Greene, J. P., Peterson, P. E., Du, J. 1998. "School Choice in Milwaukee: A Randomize Experiment." In P.E. Peterson and Bryan Hassel *Learning from School Choice.* Washington DC, Brookings Institution.

Goldhaber D. D., Brewer, D. J., Eide, E. R., and Rees, D. I. 1999. "Testing for Sample Selection in the Milwaukee School Choice Experiment." *Economics of Education Review.* Forthcoming.

Goldhaber D. D. 1996. "Public and Private Schools: Is School Choice an Answer to the Productivity Problem?" *Economics of Education Review.* 15, 93–109.

Goldhaber, D. D. and Brewer, D. J. 1997. "Why Don't Schools and Teachers Seem to Matter? Estimating the Impact of Unobservables on Educational Productivity." *Journal of Human Resources.* 32, 505–523.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica.* 47, 153–161.

Rouse, C. E. 1997. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Voucher Program." *Quarterly Journal of Economics.* Forthcoming.

Witte, J. F. 1997. "Achievement Effects of the Milwaukee Voucher Program." Paper presented at the annual meeting of the American Economics Association.

Witte, J. F., Thorn, C. A. 1996. "Who Chooses? Voucher and Interdistrict MPCPs in Milwaukee." *American Journal of Education.* 104, 186–217.

Acknowledgments

This report was prepared while the author, Dan Goldhaber, was an employee at the CNA Corporation.

The authors would like to thank David Figlio and Duncan Chaplin for their helpful comments on earlier versions of this paper. All remaining errors in this paper are ours. A more technical version of this paper is forthcoming in *Economics of Education Review.*

School-Level Resource Allocation in the Chicago Public Schools

Ross Rubenstein

Georgia State University

About the Author

Ross Rubenstein is an Assistant Professor of Public Administration and Urban Studies in the School of Policy Studies and College of Education at Georgia State University. His research focuses on educational policy and finance, including funding equity, performance, measurement, and the link between school resource allocation and performance. His dissertation, which was awarded the Jean Flanigan Outstanding Dissertation Award by the American Education Finance Association, examined these issues in the Chicago Public Schools. Dr. Rubenstein is co-author of book chapters on education finance and the measurement of school efficiency, and has published research in the *Journal of Education Finance* and *Public Budgeting and Finance*. He heads the Education Finance Studies Project at Georgia State University and currently serves as a staff member for Georgia Governor Roy Barnes' Education Reform Study Commission. Dr. Rubenstein received his Ph.D. from the Robert F. Wagner Graduate School of Public Service at New York University in 1997.

School-Level Resource Allocation in the Chicago Public Schools

Ross Rubenstein

Georgia State University

Introduction

In recent years, many public and private organizations have sought to improve responsiveness and productivity by shifting decision-making responsibilities away from centralized bureaucracies to "front-line" workers. This shift in control has strongly taken root in public education, which has seen a widespread movement among districts to adopt school-based management and budgeting systems. Although important details may vary from one district to another, school-based budgeting is intended to provide school-level personnel with much greater discretion to set spending priorities. While many districts around the country have adopted some form of school-based budgeting, little is known about the ways in which schools allocate their resources under a decentralized system. This study helps to fill this gap by examining resource allocation at the school level in the Chicago Public Schools (CPS).

Until recently, most studies of educational resource allocation have focused on districts as the unit of analysis. Technological advances and improved data availability make it possible for researchers to begin considering the school as the primary unit of analysis. As Berne and Stiefel (1994) point out, these technological changes, combined with "a growing belief that the most critical activities are closest to the child—at the school or program level," have shifted the focus of much resource allocation research to the school level. As decentralization increasingly gives schools control over budgeting, meaningful differences in resource allocation patterns among schools are likely to emerge. Moreover, econometric studies exploring educational production functions have shown that the use of micro-level data, as compared to more aggregated data, is less likely to cause specification errors leading to biased results (Hanushek et al. 1996).

In the "traditional" school governance and finance model, in which a central district office allocates funds to schools for prescribed uses, one might expect to find little intra-district variation in the allocation of money for various purposes. However, in districts practicing some form of school-based budgeting, real differences may begin to emerge in the way schools choose to target their resources to improve student performance. School-level analyses in these districts could help to determine which types of resource allocation patterns appear to be most effective for promoting specific outcomes among diverse student populations.

This paper examines the ways in which Chicago schools allocate their resources by analyzing line-item budgeted expenditures for the 1994–95 school year. It explores the relationship between spending patterns and student achievement by comparing resource allocation patterns across groups of "higher performing" and "lower performing" schools. The next section provides context for

the study by describing the reform efforts in the Chicago Public Schools. The third section reviews the data and analytic methods used in the study. The fourth section describes the results of the study, and the final section presents implications and conclusions.

Decentralization in the Chicago Public Schools

Chicago is an excellent place to conduct research on school-level spending because of the extensive decentralization efforts that began in 1989, following the passage of the Chicago School Reform Act of 1988 (P.A. 85-1418). The Act gave Local School Councils (LSCs) primary responsibility for management and budgeting decisions within their respective schools (Bryk and Sebring 1991). As the third largest district in the nation and the largest to adopt a decentralized budgeting process system-wide, the CPS includes over 550 schools making resource allocation decisions. The large number of sites facilitates statistical analyses of differences in expenditure patterns across schools.

The primary focus of the Chicago School Reform Act was "to make the individual local school the essential unit for educational governance and improvement" by shifting planning, management, and budgeting responsibilities to the school level (P.A. 85-1418). Along with this vertical decentralization of responsibility, the Act also horizontally decentralized school management through the establishment of elected LSCs at each school site. The LSCs, consisting of six parents, two community representatives, two teachers and a school principal,[1] were given the authority to hire and evaluate principals and to develop school improvement plans and school-level budgets.

Along with these additional responsibilities, schools received substantial new discretionary funding from Illinois's State Chapter 1 program. Illinois's Chapter 1 program, like the federal Title I program, distributes "compensatory" aid to schools (through their districts) based on enrollments of students from low-income families. The Chicago School Reform Act stipulated that these funds should be spent primarily on school-level instructional expenditures, and that resource allocation decisions should increasingly be made by schools and their LSCs (Bryk et al. 1993).

Despite these provisions of the Reform Act, continuing budget crises initially limited the LSCs' discretion. For the first several years of reform, Chapter 1 funds were used primarily to keep schools' base programs intact, effectively limiting the LSCs' budget options (Hess 1994). By the 1993–94 school year, state Chapter 1 funds—which amounted to $491,000 in the average elementary school and $849,000 in the average high school—were no longer needed to replace funding from other sources (Rosenkranz 1994). The 1994–95 school year provided one of the first opportunities for LSCs to make resource allocation decisions without many of the external constraints imposed by these budget cuts.

Not surprisingly, budget decisions have become an important focus of decision-making within individual schools.

Not surprisingly, budget decisions have become an important focus of decision-making within individual schools. Easton and Storey (1994) found that budgeting and finance issues were the third most common topics of discussion at LSC meetings, generating the highest participation among LSC members. Case studies developed by the Chicago Urban League (1995) demonstrate the variety of ways schools have chosen to allocate their Chapter 1 funds. While most schools in the sample added additional teachers in an effort to reduce class sizes, funds were also used to provide drug abuse prevention training for students and parents, to purchase computers and other equipment, to provide music and art programs, and to hire a variety of additional staff, including teaching assistants, school aides, and clerical and security personnel.

Data and Methods

Data Sources

The analyses presented below combine financial, demographic, and achievement data for all public elementary

[1] A student member is included in high schools.

BEST COPY AVAILABLE

and high schools in Chicago.[2] Financial data come from the CPS budget office,[3] while output and demographic data are from the Illinois State Board of Education's 1994–95 School Report Cards. Test score data comprise school-wide mean results on the Illinois Goals Assessment Program (IGAP), a series of state-administered tests given annually in selected grades in reading, mathematics, writing, social studies, and science.[4]

The financial data include only resources budgeted and spent directly at school sites, accounting for approximately 72 percent of total district spending ($2.1 billion out of a total district budget of approximately $2.9 billion). The district reports a variety of codes and descriptors for each line-item, facilitating aggregation into descriptive categories. For this analysis, all school-level budgeted expenditures are aggregated into functional categories: instruction (expenditures associated with direct instruction of students in classrooms, including teacher salaries and benefits); instructional support (expenditures for providing teacher, student and program support, including professional development and guidance, health, library, and media services); administration (expenditures for principal and assistant principal salaries, as well as attendance and security services); and operations (noninstructional expenditures associated with maintenance of the school building and lunchroom services).[5]

The national trend toward decentralization may actually impede progress toward more uniform school-level data collection because schools and districts are likely to resist additional state reporting requirements.

While school-level data hold considerable promise for improving the analysis of school effectiveness and efficiency, the data also present a host of potential difficulties that analysts must consider when constructing data sets and analyzing results (Berne et al. 1997). For example, school-level data may be more likely than district-level data to include complex codes and definitions that make aggregation into analytic categories difficult (Berne and Stiefel 1994). The multiple roles played by many school personnel can also complicate reporting and aggregation since staff duties may span several program and functional areas (Cohen 1997). The national trend toward decentralization may actually impede progress toward more uniform school-level data collection because schools and districts are likely to resist additional state reporting requirements (Clark 1998). Despite these potential difficulties, the uncommon level of detail and information provided for each line-item expenditure in the Chicago budget data set increases the likelihood of accurate and consistent aggregation into functional categories.[6]

Relative School Performance

To identify higher and lower performing schools for subsequent analyses of resource allocation patterns, *adjusted performance measures* are constructed for each school. Ordinary Least Squares (OLS) regression analysis is used to predict each school's mean student performance on several IGAP tests based upon a variety of factors outside the control of individual schools. The approach uses an educational production function to estimate the marginal effects on school outputs (IGAP scores) of various "uncontrollable" school inputs. Comparison of each school's predicted performance with its actual performance provides an indication of which schools may be producing better- or worse-than-expected academic performance.[7]

The regression equations use five independent variables to explain variations in IGAP scores. Four elementary school equations and five high school equations are esti-

[2] Middle schools are uncommon in Chicago. Most elementary schools include grades K–8.

[3] The budget data used here were reported by the CPS budget office and supplied to the author by the Chicago Panel on School Policy.

[4] Students take the reading, mathematics, and writing examinations in 3rd, 6th, 8th, and 10th grades. Students take the social studies and science examinations in 4th, 7th, and 11th grades. Each test is scored on a scale from 0 to 500, with a score of 250 representing average performance. The writing test is scored on a scale from 6 to 32.

[5] Two additional categories for community services (primarily expenditures for community and parent outreach activities) and other (expenditures for debt service and other miscellaneous items) together comprise approximately three percent of expenditures.

[6] A detailed description of the aggregation procedures is available from the author.

[7] For a more detailed description of issues in constructing and using adjusted performance measures, see Stiefel et al. (forthcoming) and Ladd and Clotfelter (1996).

BEST COPY AVAILABLE

mated, with each equation using the same set of independent variables and a different IGAP test as dependent variable. The equations are specified as:

$$T_i = \alpha + \beta_1 LOWINC + \beta_2 LEP + \beta_3 MOBILTY + \beta_4 ENROLL + \beta_5 PARINV + \varepsilon_i$$

such that $i = 1,2,...N$,

where T_i is a mean IGAP test score for school i, N is the number of schools in the data set, α is the intercept of the equation, *LOWINC* is the percentage of students from low-income households,[8] *LEP* is the percentage of students with limited English proficiency, *MOBILTY* is a measure of student mobility,[9] *ENROLL* is total school enrollment, *PARINV* is a measure of parent involvement,[10] and ε_i is an error term with the usual properties.

The adjusted performance measures provide an alternative to the use of raw test scores as indictors of school performance. It is well known that schools serving student populations that are very mobile,[11] are predominantly from low-income families, or have limited English proficiency will tend to have lower test scores than equally effective schools serving different student populations. The equations also include school enrollment as an independent variable because recent research suggests that larger schools may be less effective than smaller schools at fostering higher levels of student achievement, particularly in schools with large proportions of students from low-income families (Fowler and Walberg 1991; Lee and Smith 1997). Parent involvement is included because research indicates that schools with higher levels of parental involvement may have advantages in promoting student achievement (Henderson 1994).[12]

Table 1 displays the results of the four elementary school equations, listing dependent variables across the top row and independent variables down the left-hand column. As shown by the R^2, each equation explains between 47 percent (for third grade mathematics scores) and 58 percent (for sixth grade reading scores) of the variation in the dependent variable. In each equation, the percentage of low-income students, the percentage of students with limited English proficiency, student mobility, and school size are statistically significant. Parent involvement is significant in each equation except for the equation using third grade mathematics scores as the dependent variable. All independent variables, with the exception of the percentage of students with limited English proficiency, have the expected signs (for example, higher student mobility, higher poverty, and larger schools are associated with lower performance).

The coefficient on the LEP variable has a positive sign, which is unexpected and is particularly surprising for the reading tests. This relationship could be spurious if, for example, the result is caused by the high negative correlation between the percentage of students with limited English proficiency and the percentage of students in each school who take the IGAP tests (Pearson correlation of approximately -0.87). Therefore, the results likely reflect the more limited sample of students who take the tests in schools with high LEP populations rather than any positive effect on scores caused by limited English proficiency. Alternatively, if students for whom English is not the native language are extremely successful at overcoming this barrier, the results could reflect a real relationship between the variables.

Table 2 presents the results of the high school equations. Since high school students take reading and mathematics tests only in 10th grade, social studies, science, and writing scores are also used as dependent variables. As in the elementary school analyses, low income and mobility are consistently significant. The coefficients for the percentage of students with limited English proficiency are again generally positive (except in the equation for

[8] Low-income students are defined as those from families receiving public aid, living in institutions for neglected or delinquent children, being supported in foster homes with public funds, or eligible for free or reduced-price lunches.

[9] Student mobility is measured as the number of students moving into and out of the school during the year, divided by the October enrollment.

[10] Parent involvement is measured as the percentage of students whose family made at least one contact with the school during the year.

[11] Labeling student mobility as an uncontrollable factor can be debated. Kerbow (1995) estimates that up to 40 percent of student transfers in the CPS are due to factors, such as perceived safety problems, associated with the schools themselves. He cautions, however, that assessing school performance without adjusting for mobility will result in misleading findings.

[12] The inclusion of parent involvement as an explanatory variable assumes that, for the most part, schools cannot alter the level of parent participation, although this assumption may not always hold true.

Table 1.—Results of school-level regression analyses to predict IGAP test scores: Chicago elementary schools, 1994–95

	3rd grade reading	3rd grade mathematics	6th grade reading	6th grade mathematics
Constant	289.90***	354.14***	321.67***	319.06***
	(15.25)	(18.58)	(15.47)	(14.58)
Student mobility	-0.675***	-0.802***	-0.750***	-0.621***
	(0.121)	(0.147)	(0.135)	(0.127)
Percent from low-income families	-1.45***	-1.55***	-1.69***	-1.48***
	(0.100)	(0.122)	(.100)	(0.094)
Percent with limited English proficiency	0.835***	0.897***	0.560***	0.619***
	(0.110)	(0.134)	(0.111)	(0.104)
Total school enrollment	-0.036***	-0.043***	-0.024***	-0.025***
	(0.007)	(0.009)	(0.007)	(0.007)
Parent involvement	0.428***	0.197	0.337***	0.342***
	(0.134)	(0.163)	(0.139)	(0.131)
R^2	0.541	0.479	0.587	0.553
F	100.37***	78.36***	118.71***	103.28***
N	432	432	423	423

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

NOTE: Standard errors in parentheses.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

10th grade reading), but no longer significant, possibly because the correlation between the percentage of LEP students and the percentage of students taking the exams is much lower for high schools (-0.37) than for elementary schools.

A preferable specification of these equations might attempt to control for student ability by using a "value-added" measure of achievement as the dependent variable (such as the change in scores from one year to the next), or by including previous test scores as control variables (Hanushek 1989). Because the State of Illinois gives IGAP exams in different subject areas in different years (for example, reading and mathematics in 10th grade followed by science and social studies in 11th grade), no comparable IGAP data are available for cohorts of students from one year to the next.[13]

[13] For this study, many models were tested, including several that approximated a "value-added" approach by including third grade mathematics and reading scores as independent variables in equations using fourth grade science and social studies scores as independent variables. While coefficients and R^2 values differed, all models produced very similar lists of higher and lower performing schools. Moreover, high student mobility confounds attempts to measure changes in achievement since, for example, a fourth grade class may be very different from the third grade class of the previous year. The problem is even more pronounced between multiple grades. Only 3.8 percent of CPS retain at least 75 percent of the same students between third and sixth grade, when the IGAP reading and math examinations are administered (Kerbow 1995). Therefore, a "value-added" model using comparable IGAP scores across years would capture only a small number of the same students.

Table 2.—Results of school-level regression analyses to predict IGAP test scores: Chicago high schools, 1994–95

	10th Grade reading	10th Grade mathematics	11th Grade science	11th Grade social studies	10th Grade writing
Constant	245.81***	253.47***	265.32***	267.49***	27.34***
	(29.97)	(30.05)	(23.77)	(28.62)	(1.20)
Student mobility	-1.44***	-1.12***	-0.694***	-1.02***	-0.056***
	(0.288)	(0.288)	(0.228)	(0.275)	(0.011)
Percent from low-income families	-0.929***	-0.885***	-0.937***	-1.21***	-0.044***
	(0.295)	(0.296)	(0.234)	(0.282)	(0.012)
Percent with limited English proficiency	-0.074	0.078	0.090	0.071	0.0097
	(0.448)	(0.449)	(0.355)	(0.427)	(0.018)
Total school enrollment	0.019***	0.019***	0.015***	0.019***	-0.0004
	(0.007)	(0.007)	(0.005)	(0.006)	(0.000)
Parent involvement	-0.085	-0.261	-0.238	-0.231	-0.0029
	(0.251)	(0.252)	(0.199)	(0.240)	(0.010)
R^2	0.610	0.533	0.545	0.602	0.598
F	17.86***	13.01***	13.64***	17.21***	16.93***
N	63	63	63	63	63

* Significant at the 10 percent level.

** Significant at the 5 percent level.

***Significant at the 1 percent level.

NOTE: Standard errors in parentheses.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

Using the regression models, standardized residuals are computed for all schools. The standardized residual is the simple residual for each school on each test—defined as the difference between the observed value and the expected value—divided by the standard error of the simple residual. The measure provides a method to compare a school's *actual* performance to its *predicted* performance, given the observed characteristics of the school and its student body. Standardizing the residuals permits comparison and aggregation across multiple dependent variables measured on different scales. To account for school performance on multiple output measures, the standardized residuals are summed to produce aggregate residuals for each school, and these aggregates are used to identify groups of higher and lower performing schools.[14]

[14] The process of adding the residuals to produce an aggregate value has some drawbacks. One of the primary dangers is that use of summed residuals will make schools with very different patterns of performance on the individual tests appear to be very similar. For example, extremely high residuals on one test can cancel out extremely low residuals on another. Averaging the standardized residuals would not avoid this problem. The standardized residuals for each school are, for the most part, consistent across subject areas, indicating that schools with very high performance in one area are unlikely to have very low performance in other areas.

Resource Allocation Patterns across Higher and Lower Performing Schools

To highlight differences between higher and lower performing schools, the resource allocation analyses presented here include only "outliers"—schools with summed standardized residuals larger than plus or minus two. These outlier groups consist of 98 higher performing elementary schools, 95 lower performing elementary schools, 18 higher performing high schools and 22 lower performing high schools.

Table 3 presents the results of two-tailed t-tests comparing mean characteristics and spending patterns across the higher and lower performing elementary schools. The first four rows display the independent variables included in the regression analyses. Since the equations used to choose higher and lower performing schools attempt to control for differences in these characteristics, the lack of any significant differences between the groups is expected and indicates that—as measured by the exogenous characteristics included in the equations—the groups of schools are quite similar.

The next four rows show differences in average school performance on each of the IGAP examinations used to select higher and lower performing schools. The higher performing schools, on average, score significantly higher than the lower performing schools on each test, despite having almost identical proportions of students from low-income families, students with limited proficiency in English, and students who transfer in or out of the school.

It is important to note that the terms "higher performing" and "lower performing" are used to describe the performance of schools measured relative to each other rather than against a fixed standard of achievement. The use of regression analysis to construct adjusted performance measures focuses the analysis on the average per-

Table 3.—Difference of means: Higher and lower performing elementary schools

Variable	Means: Lower performing	Means: Higher performing	Difference
Percent low income	77.32%	78.80%	1.48%
Percent limited English proficient	13.31	12.93	-0.38
Mobility	28.64	30.26	1.72
School enrollment	635	690	55
Mean IGAP 3rd grade reading	147	221	74***
Mean IGAP 3rd grade mathematics	170	258	88***
Mean IGAP 6th grade reading	157	225	68***
Mean IGAP 6th grade mathematics	180	242	62***
Total budget per pupil	$5,414	$5,328	-$87
General fund budget per pupil	2,850	3,161	311
State Chapter 1 per pupil	698	627	-71***
Federal Title I per pupil	586	299	-287***
Average teacher salary	38,333	38,419	86
Pupil teacher ratio	16.34	16.59	0.25
Percent of funds spent on instruction	61.48%	62.14%	0.66%
Percent of funds spent on inst. support	11.65	10.33	-1.42***
Percent of funds spent on administration	6.18	6.02	-0.16
Percent of funds spent on operations	18.59	19.96	1.37**

* Significant at the 10 percent level.

** Significant at the 5 percent level.

***Significant at the 1 percent level.

NOTE: $N = 98$ high performing and 95 low performing.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

Table 4.—Difference of means: Higher and lower performing high schools

Variable	Means: Lower performing	Means: Higher performing	Difference
Percent low income	68.62%	71.48%	2.86%
Percent limited English proficient	3.95	5.87	1.92
Mobility	25.58	30.23	4.65
School enrollment	1,616	1,620	4.00
Mean IGAP 10th grade reading	141	193	52***
Mean IGAP 10th grade mathematics	147	205	58***
Mean IGAP 10th grade writing	22	24	2***
Mean IGAP 11th grade science	169	214	45***
Mean IGAP 11th grade social studies	147	199	52***
Total budget per pupil	$5,452	$5,405	-$47
General fund budget per pupil	2,980	3,334	354
State Chapter 1 per pupil	527	489	-38
Federal Title I per pupil	206	131	-75
Average teacher salary	40,649	41,026	377
Pupil teacher ratio	16.53	16.72	0.19
Percent of funds spent on instruction	62.93%	62.13%	-0.80%
Percent of funds spent on inst. support	10.78	11.18	0.40
Percent of funds spent on administration	7.05	6.03	-1.02*
Percent of funds spent on operations	18.58	20.05	1.47

* Significant at the 1 percent level.

** Significant at the 5 percent level.

***Significant at the 10 percent level.

NOTE: *N* = 18 high performing and 22 low performing.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

formance of schools in the sample. Schools producing test scores at a specified level above the average (given their uncontrollable characteristics) are identified as "higher performing" than their peers. The term "higher performing" must be used with caution, however, when examining districts in which the performance of nearly all schools might be considered unacceptable. As seen in tables 3 and 4, the average performance of the higher performing schools (particularly the high schools) was below the state benchmark for average performance. The performance of these schools was still well above that of the district's lower performing schools, however.

The next four rows compare sources of per pupil funding across the two groups of elementary schools. While the difference in total funding is relatively small and not statistically significant, the means of the three other funding sources are significantly different between the higher and lower performing schools. Lower performing schools receive, on average, more State Chapter 1 funds per pupil and substantially more federal Title I funding per pupil.[15] In contrast to the distribution of these categorical funds, higher performing schools tend to receive significantly more money from the General Fund.[16]

The next two rows in table 3 examine differences in the use of teachers. The results indicate that lower performing schools tend to have fewer pupils per teacher at a slightly lower average salary, although these differences are small.

[15] At that time, Title I (Chapter 1) funds were distributed based on both student poverty and low performance.

[16] See Rubenstein (1998) for further discussion of intra-district equity in the CPS.

The remaining rows show the average proportion of total school-level funding spent in each of four major functional areas. The variables are specified as percentages of total spending (rather than absolute dollar amounts per pupil) because, as described above, funding levels may vary between higher and lower performing schools. Schools with higher overall funding levels have the ability to spend higher dollar amounts, but not higher percentages, in all functional areas.

Few differences in overall spending patterns emerge across the two groups of schools. The results show that higher performing elementary schools tend to spend a slightly higher proportion of their budgets on instruction and operations, and a slightly lower percentage on instructional support and administration. Both groups of schools allocate slightly over 60 percent of their total resources to provide classroom instruction, a finding remarkably similar to other studies of educational spending patterns (see Odden et al. 1995 for a review of resource allocation research).

Table 4 displays the results for the sample of higher and lower performing high schools. As in the elementary schools, there are no significant differences between the groups for the independent variables used in the regression equations. Again, large differences in test scores are apparent, with the group of higher performing schools scoring substantially higher than the lower performing schools on each test.

While there are some small differences in funding levels, no clear and consistent differences emerge in resource allocations by function.

Examining categorical and general funding per pupil across the high schools, the results are similar to those for elementary schools. Although the differences are not statistically significant, lower performing schools receive greater categorical funding per pupil, but fewer dollars from the General Fund. Despite these differences in funding by source, average total budgets differ by only $47 per pupil across the two groups of schools. The comparisons of average teacher salaries and pupil-teacher ratios show that, as in elementary schools, lower performing schools average slightly lower teacher salaries and slightly lower pupil/teacher ratios.

The functional spending analyses also produce results similar to those for the elementary schools, although some differences do emerge. Unlike higher performing elementary schools, higher performing high schools spend a slightly lower proportion of their resources on instruction compared with lower performing high schools, although the differences are minimal. Lower performing high schools average approximately one percentage point higher spending on administration. As in the elementary school sample, higher performing high schools tend to spend a larger share of total resources on operations.

Overall, the spending patterns are strikingly similar across higher and lower performing schools. While there are some small differences in funding levels, no clear and consistent differences emerge in resource allocations by function. The majority of General Fund allocations, however, are for fairly narrowly defined purposes and virtually all schools must provide the same basic program to students. For example, a large proportion of General Fund allocations to schools support teacher positions, with staffing patterns determined by class size limits set in the Board of Education's contract with the Chicago Teachers' Union. Therefore, the small differences in spending patterns for total resources should not be surprising. Because the Chicago School Reform Act explicitly gave local schools the responsibility for budgeting state Chapter 1 funds, larger differences may arise in resource allocation patterns for these resources.

Tables 5 and 6 display Chapter 1 allocations by function for higher and lower performing elementary and high schools. The largest differences for elementary schools are in the proportion of resources spent on instruction, with higher performing schools spending over 3 percentage points more than lower performing schools. For high schools the difference is even larger, with higher performing schools spending over 5 percentage points more on instruction. In high schools receiving average levels of Chapter 1 funding, this difference could represent almost $50,000 in actual spending. The difference between higher and lower performing schools is largest in administration, with lower performing high schools spending almost 6 percentage points more (lower performing elementary schools spend approximately two percentage points more). Spending on instructional sup-

Table 5.—Difference of means: Higher and lower performing elementary schools, state Chapter 1 spending by function

Variable		Mean	Standard deviation	Maximum	Minimum	Mean difference
Percent spent on instruction	Lower	63.5%	15.6	100%	27.5%	3.1%
	Higher	66.6%	14.2	95.8%	24.6%	
Percent spent on instructional support	Lower	17.9	11.7	43.4	0	-.2%
	Higher	18.1	11.2	53.5	0	
Percent spent on admin.	Lower	14.3	12.4	64.4	0	-2.1%
	Higher	12.2	12.1	60.3	0	
Percent spent on operations	Lower	3.3	3.4	14.0	0	-1.0%**
	Higher	2.2	3.6	30.5	0	
Per pupil spending for security services	Lower	$17.26	$24.01	$110.55	0	-$9.97***
	Higher	$7.29	$12.98	47.42	0	

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

NOTE: *N* = 98 high performing and 95 low performing.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

Table 6.—Difference of means: Higher and lower performing high schools, state Chapter 1 spending by function

Variable		Mean	Standard deviation	Maximum	Minimum	Mean difference
Percent spent on instruction	Lower	35.3%	17.2%	74.6%	8.7%	5.4%
	Higher	40.7	16.5	76.9	6.1	
Percent spent on instructional support	Lower	29.7	13.7	70.5	2.2	0.4%
	Higher	30.1	10.4	57.5	15.7	
Percent spent on admin.	Lower	24.9	13.4	46.0	0.8	-6.0%
	Higher	18.9	9.2	36.4	5.4	
Percent spent on operations	Lower	9.5	10.6	40.9	0	-0.5%
	Higher	9.0	10.7	38.7	0	
Per pupil spending for security services	Lower	$55.12	$65.75	246.54	0	-$13.29
	Higher	$41.83	$54.46	182.87	0	

* Significant at the 10 percent level.

** Significant at the 5 percent level.

***Significant at the 1 percent level.

NOTE: *N* = 18 high performing and 22 low performing.

SOURCE: Author's calculation based on data from the Illinois State Board of Education, the Chicago Public Schools, and the Chicago Panel on School Policy.

port and operations is nearly identical for both groups of schools.

The discretionary spending differences between elementary schools and high schools are also notable. As expected, elementary schools tend to spend the majority of their Chapter 1 resources (over 60 percent) on instructional expenditures, with the next largest shares going to instructional support and administration. High schools also spend the largest share of Chapter 1 funds on instruction, however, their spending is much more evenly divided across all functional areas. Higher proportional spending in areas other than instruction could, in part, reflect the different needs of high schools and high school students compared with those of elementary schools and students. For example, high schools might have greater needs for guidance staff, which would be included in the instructional support category, and for security and disciplinary staff, which would be included in the administration category. The much higher spending for instructional support is undoubtedly also due, however, to 1991–92 budget cuts that shifted many high school aide positions from schools' base funding to their discretionary funds (Goertz and Hess 1998).

[E]lementary schools... spend the majority of their Chapter 1 resources... on instructional expenditures... High schools also spend the largest share of Chapter 1 funds on instruction, however, their spending is much more evenly divided across all functional areas.

The cumulative evidence suggests that higher performing schools spend a greater share of their resources on direct instructional expenditures, but it does not explain the causality of these patterns. It is unclear whether students in higher performing schools perform better than those in low performing schools because of a greater emphasis (at least financially) on instruction, or whether higher performing schools are able to spend larger shares of their resources on instruction because their students achieve at higher levels and have fewer needs in other areas.

Alternatively, performance and spending patterns may be simultaneously determined; that is, both may be results of the school's organization and the performance of teachers and staff. Lower performing schools' higher discretionary spending on administration could reflect a less efficient use of resources by these schools, or it could indicate that lower performing schools tend to have greater needs in administrative areas such as attendance and security.[17] The bottom rows in tables 5 and 6 bear out this latter hypothesis, showing a significant difference between higher and lower performing schools in Chapter 1 spending per pupil for security services. While these data provide no indication of whether these schools have greater real needs for security to justify this higher spending, additional money spent on security services leaves less money for classroom instruction.[18]

Tables 5 and 6 also display differences in the maximum and minimum values for the Chapter 1 spending variables. Table 5 shows that at least one (lower performing) elementary school spends its entire Chapter 1 allotment on instruction, while another spends only about one-quarter in that area. Some elementary schools spend no Chapter 1 money in each of the other functional areas. No high school spends more than 77 percent of its Chapter 1 funds in a single functional area. Overall, as measured by the range and standard deviation of spending by function, it appears that high school Chapter 1 spending patterns exhibit more variation than elementary school spending patterns, with the possible exception of spending for administration. The data also demonstrate that, despite the mandate in the Chicago School Reform Act that discretionary funds be spent primarily for instructional purposes, many schools spend the funds in other areas.

Discussion

This paper examines the relationship between spending and student performance in a large school system practicing school-based budgeting. The study finds some evidence that schools in Chicago spend their available discretionary resources in diverse ways and that the differences across schools in their resource allocation patterns may be related to school performance. Specifically, schools with higher-than-predicted performance tend to allocate a larger share of their discretionary resources for

[17] The "administration" category includes all spending intended primarily to improve school safety and security.

[18] The slightly higher average Chapter 1 funding among low performing schools could mitigate this effect.

BEST COPY AVAILABLE

instructional purposes, while schools with lower-than-expected performance tend to spend more in noninstructional areas, such as security. The differences are particularly striking in high schools, with higher performing schools spending six percentage points less on administration and an almost equal amount more on instructional expenditures, compared with the lower performing schools.

The study also finds that while total spending patterns are relatively consistent across groups of schools, discretionary spending choices show more variation. The comparisons of Chapter 1 spending and total spending suggest that greater differences in resource allocation patterns may begin to emerge as schools gain greater control over their budgets; and as school staff and parents are empowered to make significant spending decisions. As long as resource allocation decisions are made only at the margins, it may be unlikely that school-based budgeting will have a large or appreciable effect on student performance. As schools in Chicago and elsewhere gain greater control over their entire pool of resources, we may begin to see larger and more meaningful differences in spending patterns across schools. These differential spending patterns might be considered the first "outputs" of decentralized school budgeting processes.

References

Berne, R., L. Stiefel, and M. Moser. 1997. "The Coming of Age of School-Level Finance," *Journal of Education Finance*, 22: 246–254.

Berne, R. and Stiefel, L. 1994. "Measuring Equity at the School Level: The Finance Perspective," *Education Evaluation and Policy Analysis*, 16: 405–421.

Bryk, A. S. and others. 1993. *A View from the Elementary Schools: The State of Reform in Chicago*. Chicago, IL: Consortium on Chicago School Research.

Bryk, A. S. and Sebring, P. A. 1991. *Achieving School Reform in Chicago: What We Need To Know*. Chicago, IL: Consortium on Chicago School Research.

Chicago School Reform Act of 1988. P.A. 85–1418. 105 ILCS 5/34–1.01.

Chicago Urban League. 1995. *Use of State Chapter 1 Funds in the Chicago Public Schools Districts 6 and 8*. Chicago: Chicago Urban League, Department of Research and Planning. ED 403344.

Clark, C. 1998. "Using School-Level Data to Explore Resources and Outcomes in Texas," *Journal of Education Finance*, 23: 374–389.

Cohen, M. C. 1997. "Issues in School Level Analysis of Education Expenditure Data," *Journal of Education Finance*, 22: 255–279.

Easton, J. Q. and Storey, S. L. 1994. "The Development of Local School Councils," *Education and Urban Society*, 26: 220–37.

Fowler, W. J., Jr., and Walberg, H. J. 1991. "School Size, Characteristics and Outcomes," *Educational Evaluation and Policy Analysis*, 13: 189–202.

Goertz, M. E. and Hess, G. A. Jr. 1998. "Process and Power in School Budgeting Across Four Large Urban School Districts," *Journal of Education Finance*, 23: 490–506.

Hanushek, E. A., Rivkin, S. G. and Taylor, L. L. 1996. "Aggregation and the Estimated Effects of School Resources," *The Review of Economics and Statistics*, 78: 611–627.

Hanushek, E. A. 1989. "The Impact of Differential Expenditures on School Performance," *Educational Researcher*, 18: 45–51.

Henderson, A. 1994. *A New Generation of Evidence: The Family is Critical to Student Achievement*. Washington, DC: National Committee for Citizens in Education.

Hess, G. A., Jr. 1994. "School Based Management as a Vehicle for School Reform," *Education and Urban Society*, 26: 3–17.

Kerbow, D. 1995. *Pervasive Student Mobility: A Moving Target for School Improvement*. Chicago, IL: Chicago Panel on School Policy.

Ladd, H. F. and Clotfelter, C. T. 1996. "Recognizing and Rewarding Success in Public Schools," *Holding Schools Accountable: Performance-Based Reform in Education,* H.F. Ladd, (ed.), Washington, DC: Brookings Institution: 3–63.

Lee, V. E. and Smith, J. B. 1997. "High School Size: Which Works Best and for Whom?" *Educational Evaluation and Policy Analysis,* 19: 205–227.

Odden, A., Monk, D., Nakib, Y., and Picus, L. 1995. "The Story of the Education Dollar: No Academy Awards and No Fiscal Smoking Guns," *Phi Delta Kappan,* 77: 161–68.

Rosenkranz, T. 1994. "Reallocating Resources: Discretionary Funds Provide Engine for Change," *Education and Urban Society,* 26: 264–284.

Rubenstein, R. 1998. "Resource Equity in the Chicago Public Schools: A School-Level Approach," *Journal of Education Finance,* 23: 468–489.

Stiefel, L., Rubenstein, R., and Schwartz, A. E. (forthcoming). "Using Adjusted Performance Measures to Evaluate Resource Use," *Public Budgeting and Finance.*

The Productivity of School Finance Equalization: An Analysis Using Hierarchical Linear Modeling

Patrick Galvin

Hal Robins

Karen Callahan

University of Utah

About the Authors

Patrick Galvin is Associate Professor, Department of Educational Leadership and Policy, and co-director of the Utah Education Policy Center at the University of Utah. His research interests focus on economics of school organization particularly as related to issues of school finance and educational productivity.

Hal Robins is currently employed as Senior Research Analyst with the Utah Education Policy Center, University of Utah. His research and policy work focus on school finance issues. Dr. Robins was previously employed with the Utah State Office of Education as a school-finance specialist. His considerable experience and expertise with Utah's school finance formulae and the legislative process contributes significantly to his work at the Utah Education Policy Center.

Karen Callahan is a recent graduate from Northwestern University, where she graduated with a degree in educational policy analysis. During the summer of 1998, Ms. Callahan worked as a summer-intern with the Utah Education Policy Center, University of Utah. She presented worked developed from that experience at the 1999 American Education Finance Association in Seattle, Washington.

BEST COPY AVAILABLE

The Productivity of School Finance Equalization: An Analysis Using Hierarchical Linear Modeling

Patrick Galvin

Hal Robins

Karen Callahan

University of Utah

Does Money Matter: School Finance Equalization and Educational Performance

The debate about the relationship between educational resources and levels of educational achievement has a long, inconclusive history (Picus 1995). In this paper, we tie productivity to traditional school finance equity issues by asking, What are the productive characteristics of school finance equalization? In other words, are the efforts made by legislators to differentiate funding by measures of school district need associated with variations in measures of performance?

The motivation for the paper is predicated on the assumption that the underlying goal of school finance equalization is to promote educational achievement. Legislators, the public, and educators are rightfully concerned about the fair distribution of available educational resources. The goal of these efforts, however, is to equalize educational opportunities and outcomes, not inputs per se.

A second purpose underlying this study is grounded in the methodology by which we examine the above-mentioned questions. We begin with the premise that the organization of schools is hierarchically nested: students are nested within classrooms, classrooms within schools, schools within school districts, etc. This is to say that two otherwise identical schools may perform differently based on the social, organizational, and fiscal environment in which they operate. In this study, we apply Hierarchical Linear Regression Models (HLM) to control for differences in school district characteristics (contextual factors) (Bryk and Raudenbush 1992). Indeed, we describe the school as operating in a resource environment, which we defined as a school district (contextual) variable.

What we find is intriguing evidence that equalization of fiscal resources is masked by an interaction of school and district socio-economic status (SES) effects. The relationship between variations in resources and their effect on school-level measures of performance (our measure of educational productivity) appears to be dependent upon the socio-economic characteristics of the district in which the school operates. Ignoring these contextual factors raises both conceptual and statistical matters that may account for some of the confounding results associated with previous studies on educational productivity.

BEST COPY AVAILABLE

Background and Relevance of the Research Problem

This report is an outgrowth from an earlier study, which examined school finance equity issues in the state of Utah over a 10-year period.[1] Using the traditional Berne and Stiefel (1984) equity design, we found evidence that both measures of horizontal and vertical equity had improved over time in Utah. One of the many trend graphs we generated to substantiate this point is included in the appendix of this paper.

Our initial reaction to this finding was that it represented good news for a state that had experienced significant financial, demographic, and policy change during the period of our study. Upon more reflection, however, it seemed to us that the findings did not completely address the concern legislators and taxpayers have regarding the topic of school finance. Certainly, the public were concerned about the fair distribution of resources allocated for education but they were also concerned about their effective use. This very point underlies Firestone, Goertz, and Natriello's (1998) study of the implementation of the Quality of Education Act (QEA) in New Jersey. The QEA redistributed significant amounts of public money to equalize funding for the state's urban-poor school districts. While political support following the New Jersey State Supreme Court decision in Abbott v. Burke, 119 N.J. 287 (1990) (Abbott II) was sufficient to pass the QEA, it withered when questions were raised about how effective such investments would be in improving educational outcomes. The authors provide extensive documentation that the QEA money used for equalization largely contributed to improving the educational opportunities of the students targeted by the court case.

While political support following the...decision in Abbott v. Burke...was . sufficient to pass the QEA, it withered when questions were raised about how effective such investments would be in improving educational outcomes.

The issues that underlie the QEA in New Jersey reflect the concerns we had about the equity study in Utah. We characterized this point by suggesting that a policy gap existed between the conventional focus of school finance studies on the distribution of educational inputs (i.e., resources) and the public's concern about the productive use of those resources. This gap in policy goals between equity and productivity underlies the focus of the study reported in this paper. We wanted to explore the productive effects of equalization, because legislative funds are, in our opinion, allocated to promote learning as well as to achieve fiscal equity.

Approach to the Study

Our approach to this topic follows the logic presented by Bryk and Raudenbush (1992) in their book, *Hierarchical Linear Models: Applications and Data Analysis Methods*. These authors argue that if one regresses measures of achievement on SES, a comparison of the intercept and slope provides a means for judging effectiveness and equity (p. 10). This approach is not unfamiliar to research on educational productivity, which relies on a single regression line through all of the schools included in a study. An illustration of such an assessment is presented in figure 1, using data from Utah's elementary schools, in which school-level measures of performance are regressed on school-level measures of SES.

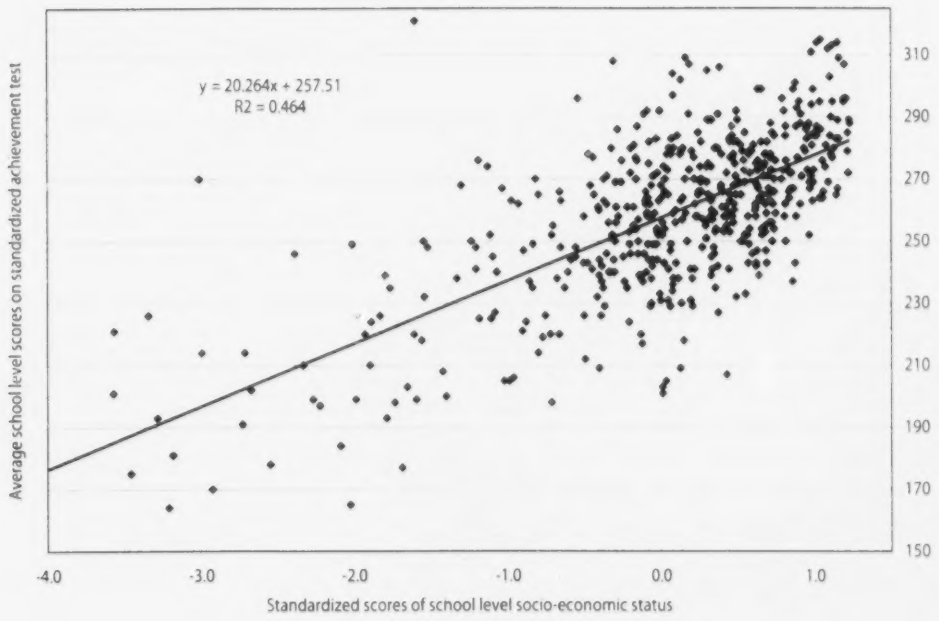
Before discussing the point of including figure 1 in this paper, it should be noted that the measure of a school's SES is the average of three standardized measures: 1) percentage of free and reduced-price lunches; 2) the percentage of low-income students; and 3) percentage of AFDC students. These three scores were reversed, which resulted in higher scores indicating higher levels of SES. These scores were standardized and then averaged for each school. These *z*-scored measures are centered around zero (0.019) with a standard deviation of 1 (0.98). Cronbach's Alpha was 0.86, which suggests a highly reliable scale.

School measures of achievement are the product of the Utah's Statewide Assessment Test (SAT), a nationally normed performance test, given to all 5th-, 8th-, and 11th-grade students. There are five batteries in this test

[1] In 1989 and 1990, the Utah State Legislature enacted a number of changes in the plan by which public schools are financed. Additionally, during the 1990s, significant changes in local property taxes and the assessed valuation of property occurred within the state. Thus, our primary motive in the equity study was to assess how changes in state finance policies, property taxes, and property values had affected school finance equity within Utah.

BEST COPY AVAILABLE

Figure 1.—Ordinary least squares (OLS) regression of socio-economic status (SES) on achievement of Utah elementary 5th-graders, 1994–95 (502 schools)



SOURCE: Graph calculated by author from data collected by the Utah State Office of Education.

(mathematics, English, reading, science, and social studies). We have used the "Basic Battery," which samples the five areas for a general assessment. The data used for these regressions were raw scores, number of correct answers, not percentile ranks, which would be inappropriate because the intervals for the percentile ranks vary.

Thus, a school's score is the average number of correct answers for the students taking the test; the average test group was 134 students, ranging from a minimum of 20 to a maximum of 558. Scores range from a minimum of 156 to a maximum of 350, with the average at 258 correct answers.

The results of this regression suggest that the composite measure of student SES by school accounted for about 46 percent of the variation in performance scores. The scatter plot shows a positive relationship between SES and achievement, with a single regression line through the data.

If the conditions associated with these data were experimental (i.e., students and teachers were randomly assigned within a controlled environment), a comparison of regression results by year would enable one to make judgments about the effectiveness and equity of the system over time. That is, if the measure of the centered intercept increased from year to year, then one could argue that the school system was performing more effectively, controlling for the SES of the students within schools. Similarly, if the slope of the regression line flattened, one could argue that school system was producing a more equitable distribution of outcomes over time. Indeed, this discussion suggests the obvious point that at least two products can define productivity in education: average levels of achievement, equalization of achievement, or both.

One of the assumptions associated with Ordinary Least Squares (OLS) regression is that the observations are independent of one another and that the error terms are

randomly fixed, that is, $r_i \sim N(0,\sigma^2)$. This is to say that there is homogeneity of variance within the sample, and that the variance is not systematically structured. The problem with this assumption is that it is explicitly not true. Schools are grouped within fiscal and administrative clusters known as school districts. The significance of recognizing differences in organizational context can be clearly illustrated by running OLS regressions for schools within their respective districts [the first step in Burstein's (1980) Slopes as Outcomes Model]. Figure 2 displays a sample of the results for such a series of regressions. It should be noted that these districts were hand picked to illustrate the point discussed below. Moreover, as a point of clarification, all of the 27 districts included in this analysis had at least five 5th- and 8th-grade school classes reporting achievement data (the average was 18 schools per district). Districts with fewer than five cases created outliers that made for an unstable analysis and were eliminated.[2]

Some of the resulting coefficients in figure 2 are positively sloped for schools within districts, suggesting that school performance within those districts is positively associated with increased measures of SES. In other districts, the SES-achievement slope is flat or negatively sloped, suggesting that the relationship between SES and achievement is not very strong; the reverse of what we normally expect.

The failure of OLS to properly handle error terms between organizational levels is one of the fundamental justifications for using HLM regression techniques instead of OLS.

OLS regression, conceptually and statistically, fails to account for this dependence within the structure of school organization data. Additionally, OLS (and Slopes as Outcomes) fails to properly partition the variance between districts and schools. The failure of OLS to properly handle error terms between organizational levels is one of the fundamental justifications for using HLM regression techniques instead of OLS.

Conceptually, OLS regression fails to recognize the hierarchical structure of school organization. One obvious aspect of this structure is related to school funding. State legislatures allocate funds to a fiscal agent, known as a school district. Tracking resources once they reach the school district is troublesome for numerous reasons: budgeting is largely a set of balance sheets not a resource flow chart; education operates with considerable spill-over effects and jointedness in production making it difficult to properly account for resources (Monk 1990). Thus, rather than asserting that each school, or student, operates with X number of dollars per unit, it seems more appropriate to state the obvious: that students and schools operate within a resource environment, which we call a school district.

Fiscal resources available to support the production and delivery of educational services vary significantly among Utah's school districts. For example, per pupil revenues for instruction in the 1994–95 school year average $2,331 per pupil; but these figures ranged from $1,946 to a maximum of $3,372 per pupil. Recognition of these different fiscal and administrative environments helps explain why two otherwise identical schools located in different school districts may perform differently on statewide performance tests.

Data Structure, Research Design, and Questions

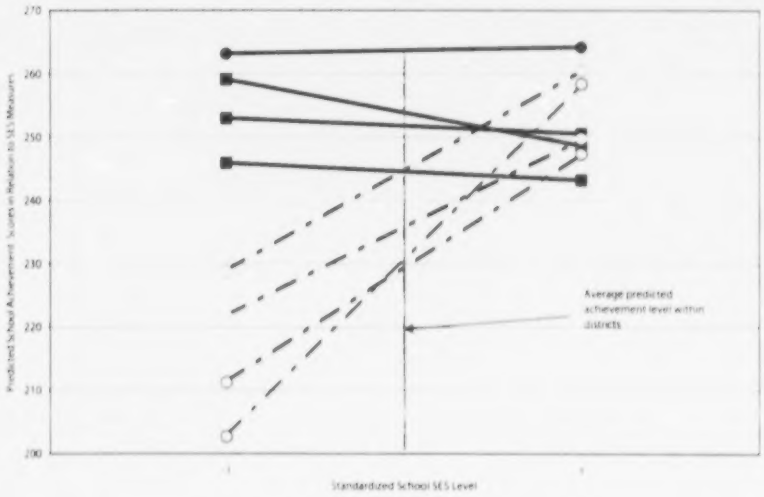
The above discussion highlights the conceptual framework guiding our inquiry. It also identifies some key elements of our data: revenue figures, student counts, measures of SES, and statewide test scores. Our methodological approach to the study relies on the use of HLM regression techniques. This method requires two data sets, appropriately designated level 1 and level 2, school and district variables, respectively.

Level 1 data included school-level variables. We collected data on a variety of measures but two figure prominently in the study: school-level measures of SES and school-level measures of academic performance. Both of these variables were introduced in the discussion earlier in this paper. The total number of schools included in the study was 502, which included all 5th- and 8th-grade classes

[2] As an example of the point justifying this decision, several of the smaller school districts eliminated from the study had regression coefficients of more than 200 compared with the overall mean coefficient of about 25.

BEST COPY AVAILABLE

Figure 2.—Varying slopes as outcomes (selected regression coefficients by district)



SOURCE: Graph calculated by author from data collected by the Utah State Office of Education.

reporting SAT scores. Both grade levels use a similar scale for their test results and including both grades increases the number of cases per district for the analysis. To ensure reliability of test scores, schools reporting fewer than 10 students with SAT scores were deleted from the data set. This eliminated only 34 schools.

The level 2 variables represented district-level data. Two measures of revenues are used in the analysis. The first is simply the sum allocated to a district divided by the number of students counted within the district. These figures are designated within the study as revenues/pupil (REV/PUP). The second measure divides the state's allocation to districts by the count of Weighted Pupil Units (WPU). These measures will be discussed below.

One critical part of this research is defining a set of measures by which one compares an allocation reflecting district need relative to an allocation that does not. In an experimental design this would not be a problem; we would compare the performance of students in a school funded without equalization with that for funding with equalization money. Of course, this experiment is not going to happen so we have to approach the topic in a post hoc fashion.

We start by noting that, in the simplest scheme of things, there are two levels of equalization in Utah's school finance formula. First, the state's school finance formula equalizes its foundation grant. Second, through a variety of formulas and program initiatives, the special needs of school districts are identified and compensated. The foundation grant accounts for about 40 percent of the state's budget for public education; the other 60 percent is allocated for special needs. In this study it is this second tier of equalization that interests us.

As noted above, in a controlled research environment we would vary the level of funding and compare the resulting levels of performance. What we have, however, is both a total allocation to school districts, for their foundation and special needs, and the resulting test scores. What does vary, and is the crux of this study, is the count of allocation units the state uses for funding equalization—the WPU. Each student within a school district receives one WPU for the basic foundation grant. Additional money for special education, vocational students, and other programs is distributed by the calculation of WPUs. This is to say that a district receiving a foundation grant for 1,000 students (WPUs) may receive an additional allocation for 500 WPUs to compensate for

its special needs. Thus, the WPU standardizes the state allocations relative to measures of district need.

Our effort to examine the productivity of equalization is organized around the difference in student head counts and the count of WPUs. If the value of the WPU is given at $2,000, then the above hypothetical district would receive as its operating budget $3,000,000. Relative to head counts, the revenues available are $3,000/pupil. Relative to WPUs, however, the revenues are only $2,000/WPU. The difference in the values reflects the effort to standardize differences in need among the school districts. One indication of how significantly these calculations vary is the correlation between the two sets of figures among Utah's 40 school districts: $r = 0.38$. We use the differences in these two sets of figures in our efforts to assess the productivity of equalization.

An additional level 2 variable is the average SES of the district. This statistic provides the means for assessing the socio-economic context of schools. This is to say, it is one thing to operate as a high-need school within a relatively wealthy environment and quite another to operate as a high-need school in a relatively poor environment. These resource and peer effects, while common to the analysis and discussion of student performance, are relevant to an assessment of school performance as well. District SES was calculated by averaging district z-scores for the percentage of students not identified as low-income with the assessed valuation per pupil; these measures ranged from 1.5 to -5.2 with a mean of 0.011.

These resource and peer effects, while common to the analysis and discussion of student performance, are relevant to an assessment of school performance as well.

The HLM Model and the Question of Centering Variables

One of the primary admonitions by scholars of multilevel models is to keep the number of parameters within the model to a minimum. Kreft and de Leeuw (1998) note that multilevel analyses provide more realistic models with greater statistic accuracy, but at a price of greater instability in the results. Like Bryk and Raudenbush (1992), these authors recommend keeping the model small. We have followed that advice.

Additionally, we should note that we have not centered the level 1predictors (e.g., SCHLSES). Centering does not rescale variables monochronomically and, hence, alters the data in substantive ways. We were convinced by Kreft and de Leeuw's (1998) argument that centering raised significant questions about the nature of the modeling. The use of raw scores for level 1 variables seemed most defensible. We did, however, center level 2 variables around the grand means, which, as Kreft and de Leeuw note, does not effect the character of the results but does make interpretation slightly easier.

The full HLM model used in this study regresses school-level SES on measures of school achievement, while controlling for variations in district levels of funding and measures of SES. The full HLM regression model, displayed in figure 3, is represented as described in the following paragraph.

With these points in place, the remainder of the paper describes the results of our exploratory study. We begin by describing the degree of dependence intraclass correlation (ICC) school achievement data has with district organization. This finding in itself is important. Next we compare the regression coefficients for school SES on performance measures from both HLM and OLS. Additionally, the analysis provides an opportunity to compare the percentage of variance explained within the HLM regressions, which is considerably less than calculated in OLS (which is subject to the problem of aggregation bias). Our third analysis examines the relationship between district-level variables and school measures of performance. The fourth analysis summarizes the whole model introduced above. We close the paper with a summary discussion. Finally, we use the deviance statistics (Chi-square) to assess whether the differences which exist between the WPU and enrollment revenue models and measures of achievement were significant.

First Analysis: Evidence of Dependency within the Data

A fundamental assumption justifying the use of HLM is the presence of dependency within the data structures

BEST COPY AVAILABLE

Figure 3—Full Hierarchial Linear Regression Model (HLM)

Level 1 Model

$Y = B0 + B1\ (SCHLSES) + R$

Level 2 Model

$B0 = G00 + G01\ (DISTSES) + G02\ (REV/PUP) + U0$

$B1 = G10 + G11\ (DISTSES) + G12\ (REV/PUP) + U1$

Where Level 1:

Y = the average achievement score for each school
B0 = the intercept for each set of schools within their respective district
B1 = the slope for each set of schools within their respective district
R = random error term

Where Level 2:

B0 = the intercept for each set of schools within their respective district
G00 = the grand mean achievement measure for all districts
G01 = average school SES-achievement intercept controlling for district SES
G02 = average school SES-achievement intercept controlling for district revenues
U0 = the unique effect of district *j* on mean achievement holding district SES and revenues constant

B1 = the SES-achievement slope for schools within their respective district
G10 = the grand mean SES-achievement slope for all schools within their respective districts
G11 = average school SES-achievement slope controlling for district SES
G12 = average school SES-achievement slope controlling for district revenues
U1 = the unique effect of school *j* on SES-achievement slope holding district SES and revenues constant

SOURCE: Author's regression model.

being examined. The most basic HLM procedure, which Bryk and Raudenbush describe as the Random One-Way ANOVA, provides a reliability statistic that characterizes "...the reliability of each school's sample mean as an estimate of its true population mean" (Bryk and Raudenbush 1992, 61). These results are displayed in table 1 for each of the 5 years under investigation.[3] Scores for these statistics range from 0 to 1; the higher the statistic the more certainty one has that the scores between schools are significantly different.

These statistics suggest that significant differences in the achievement scores among schools within Utah exist and that they are highly dependent upon district characteristics. This confirms our suspicion that matters of context are important for analyzing the school finance or productivity question. Thus, one would infer that district-level variables should account for a significant portion of the total variance in school-level measures of performance. In the following section, the within-group and between-group variance of performance is partitioned; these results are compared using HLM and OLS.

Partitioning Variance: Comparing HLM and OLS ANOVA Results

Bryk and Raudenbush (1992), as well as other multilevel regression statisticians (Goldstein 1995), note that OLS regression fails to properly account for error terms and, hence, variance that otherwise could be systematically assigned to predictor variables is accounted for as random error. This is an important statistical issue in social science models which frequently explains less than 10 percent of the variance (Bridge et al. 1979). Comparing the variance components for HLM and OLS provides evidence that, in fact, HLM does account for a greater percentage of the variance across districts than does OLS. The results of this analysis are presented in table 2.

[3] Our initial plan for the study was grandiose and included comparable analyses across all 5 years. The mechanics of producing these results is not nearly as difficult as reporting them. We have included the initial report of all 5 years as evidence that there is nothing unique about the 1990–91 school year on which focus is placed later in the paper.

Table 1.—Full Hierarchial Linear Regression Model (HLM) reliability statistic school-level achievement scores nested within district organizations 1990–91 to 1994–95

	HLM reliability statistic ANOVA
1990–91	0.745
1991–92	0.810
1992–93	0.717
1993–94	0.736
1994–95	0.675

NOTE: Schools *N*=502; districts *N*=27.

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

Using the basic statistical concepts of ANOVA in both HLM and OLS reveal remarkably different results. The OLS ANOVA indicates that only 16 to 23 percent of the total variance in scores is accounted for between districts. The One-Way Random ANOVA in HLM indicates that from 30 to 38 percent of the variance is accounted for at the district level. An increase of 13 to 15 percent of the variances is significant and represents an important finding of this research.

Analysis Two: Influence of District Variables on Performance

In this section of the report, the relationship between district-level variables and achievement is examined. We begin with district measures of SES (DISTSES) and measures of revenue (REV/PUP and REV/WPU). In this respect, the school-level model introduced in the One-Way random ANOVA remains unchanged: school-level performance scores are viewed as varying around their district mean. The district-level model is now elaborated so that level 2 predictors condition each school's mean. Our purpose is to find out if these predictors are significantly related to achievement, the direction of their coefficients, and the percentage of the between-group variance each explains. Additionally, we compare these results over a five-year period to assess trends. Finally, we specifically compare the difference between the REV/PUP and REV/WPU models, because the results provide evidence of whether efforts to promote vertical equalization are significantly and positively related to school-level measures of performance.

The initial report focuses on a single year, 1990–91, largely because of the complexity of reporting results for more than one year. The basic model used for this analysis, called the "Means as Outcomes" regression (Bryk and Raudenbush 1992), is presented in figure 4. There is no school-level (level 1) predictor for achievement in this model. The three conditioning district-level variables (level 2) are entered individually in separate models.

The results of the analyses are presented in table 3. In order to keep the presentation as focused as possible, the table includes information only about each of the predictor variables. The SES of the district (the *z*-score average of district wealth and percentage of students not identified as low-income) is significantly and positively associated with increased levels of school performance: the regression coefficient is 15.9 (p value = 0.011). Thus, for every point district SES increases, the average school-level performance score increases 15 points. A single point increase on DISTSES is equivalent to a 10 percent reduction of low-income students, or about a $200,000 increase in assessed valuation. In other words, schools operating in environments with both low incidence of poverty (high SES) and in high assessed valuation perform, on average, much higher than comparable schools

Table 2.—Comparison of variance partitioning using HLM and OLS 1990–91 to 1994–95

	ANOVA: Variance accounted for between districts		
	HLM	OLS	Difference as percent
1990–91	33.7%	19.9%	13.9%
1991–92	38.4%	23.6%	14.8%
1992–93	32.0%	17.1%	14.9%
1993–94	33.2%	18.8%	14.4%
1994–95	29.7%	16.1%	13.6%

NOTE: Schools *N*=502; districts *N*=27.

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

Figure 4.—Means as outcome Full Hierarchial Linear Regression Model (HLM)

Level 1 Model
Y = B0 + R
Level 2 Model
B0 = G00 + G01 (DISTSES) + U0

And then subsequently
B0 = G00 + G01 (REV/PUP) + U0
B0 = G00 + G01 (REV/WPU) + U0

SOURCE: Author's regression model.

Table 3.—Hierarchial Linear Regression Model (HLM) regression outcomes, 1990–91

Fixed effect	Coefficient	Standard error	T-ratio	P-value
DISTSES90, G01	15.884621	5.756734	2.759	0.011
REV/PUP90, G01	-0.025560	0.007892	-3.239	0.004
REV/WPU90, G01	-0.050963	0.022775	-2.238	0.034

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

in high-need and low-resource environments. It is difficult, however, to draw inferences from these findings since we have not yet controlled for differences in school characteristics. Nor have we controlled for differences in resource levels. We have only found that district SES is positively associated with average school performance among districts.

The second regression examines the relationship between varying levels of revenues per pupil (REV/PUP) and average school performance among districts. This revenue figure does not account for differences in district need, it only reflects the revenues available per student. The results, in the second line of table 3, indicate that increases in revenues per pupil are significantly, but negatively, related to average performance levels. Thus, as revenues increase, the average measure of performance declines.

The third regression controls for differences in district need by calculating revenues relative to the count of WPUs. Revenues per WPU (REV/WPU) are significantly, but negatively, related to school performance. To estimate the effect of this variable we note that the revenue difference per WPU is about $1,200 among Utah's school districts. Substituting these dollar figures into the regression equation, while holding school characteristics constant, suggests that about one-fifth of the variation in school scores is due to wealth differences at the district level. In other words, the performance of two otherwise identical schools would differ by about 50 points depending upon whether they were members of the wealthiest or poorest school district within the state.

Another issue relevant to these analyses is an estimation of how much of the variance these variables explain.[4] DISTSES accounted for 16.1 percent of the between group variance, while revenues per pupil (REV/PUP) accounted for almost 21 percent of the variance. The complex structure of the data and regression model preclude simply adding the explanatory power of the variables for a total estimate of the explanatory power of the regression model. The estimates of variance explained by the socio-economic status of the district (DISTSES) and then subsequently by revenues per pupil (REV/PUP) represent independent calculations that cannot be summed as an estimate of explanatory power of the regression model. Interestingly, the calculation of revenues by WPU (REV/WPU) only accounted for 9.7 percent of the between-group variance.

[4] This figure is calculated by subtracting the Tau from the DISTSES model (or the Revenue models) from the Tau for the random ANOVA District SES and dividing the difference by the Tau for the random ANOVA District SES model (see Bryk and Raudenbush 1992).

The power for DISTSES as a predictor of mean achievement increased over the years, even after the state's new formulas intended to enhance equalization were implemented.[5] By 1992–93, for example, DISTSES accounted for 42 percent of the between-group variance identified in the random ANOVA. During this same year, REV/WPU accounted for only 3.7 percent of the total between-group variance.

These findings raise at least three points of interest. First, the traditional equity study examining these same data portrayed each succeeding year as more equitable than the last (see as an illustration the trend graph in the appendix). The HLM study raises questions about such a conclusion. Rather, it appears that the changes in the state's school finance formulas strengthened the relationship between educational outcomes and district SES. Additionally, the mechanism by which vertical equity is promoted, the WPU, was not very influential in the relationship between school performance and resources. The analyses presented in this section of the paper suggest that, in Utah's case, the promotion of fiscal equity did not promote more equitable outcomes or a higher level of performance.[6]

One speculative interpretation of the data might be that the state's formula is relatively sensitive to existing accountability measures which are primarily defined in terms of traditional school finance equity statistics. Thus, while there is a lot of rhetoric about equalizing educational opportunities and outcomes, the incentive structure by which state offices of education are held accountable is defined relative to equalizing fiscal revenues, regardless of their effect on educational outcomes. Before the introduction of statistical methods such as HLM, there were relatively few ways of assessing these relationships adequately. Thus, the practice of school finance can be seen as equalizing educational inputs without much regard for their effect on educational outcomes. Obviously what we all would like is to develop legislation and policy that would promote both.

While these models provide important information about the relationship between district-level variables and average levels of performance, we have not yet controlled for differences in school-level characteristics. In the next section, we examine the relationship between school-level SES and measures of school performance. We would like to know the strength of this association and how much it varies across school districts. In other words, are school-level measures of SES a more powerful predictor of achievement in some school districts than in others? If so, then some school districts may more equitably equalize their outcomes than others. We begin with a simple random slope model, in which only school SES is included in the model; no level 2 (district) variables are included in the model.

Analysis Three: School-level SES Coefficients

In this section we analyze the relationship between measures of school socio-economic status (SCHLSES) and school measures of performance. We thus conceive of each district as having its own regression equation with an intercept and slope (Bryk and Raudenbush 1992, 67). We would like to know the average level of achievement and averages slope for schools across all the school districts, controlling for school-level measures of SES. Additionally, we are concerned about the degree of variance

[5] The table in this footnote displays the coefficients and standard errors for each of the regressions by year. Generally, they reveal a strengthening relationship between DISTSES and school performance, which is consistent with the discussion about variance explained.

	DISTSES		REV/PUP		REV/WPU	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
1990–91	15.885	5.757	-0.026	0.008	-0.051	0.023
1991–92	21.822	6.000	-0.026	0.009	-0.036	0.034
1992–93	20.391	4.349	-0.023	0.006	-0.043	0.030
1993–94	21.299	4.763	-0.018	0.008	-0.022	0.031
1994–95	19.237	4.485	-0.021	0.007	-0.003	0.028

[6] The relatively weak predictive power of the state's allocation based on need is further evidenced in the analysis of the Deviance statistics generated by the HLM output, which Bryk and Raudenbush describe as a measure of the "goodness of the model fit" (p. 54). Adding REV/WPU to the regression model did not significantly change the Deviance statistic, even though there still existed significant variance to be explained in the model. Neither did the inclusion of REV/PUP, although DISTSES did prove to be significant.

BEST COPY AVAILABLE

in these scores, as well as the correlation between slope and intercepts. In other words, is it true that high achieving districts also have large slopes (high degrees of inequity within the system)? Finally, we want to compare these results with OLS as a referent point for the discussion.

Each district's distribution of performance scores is characterized by two statistics: the intercept and the slope of the regression line. The intercept provides information about the level of performance controlling for (or removing the variance associated with) independent variables such as school size, wealth, and SES. Multilevel regression, as previously noted, also properly accounts for the variance associated with hierarchically nested data, and hence more accurately specifies standard errors in the regression. The slope provides information about the relationship between two variables, such as the relationship between school SES and measures of school performance. As noted earlier, Bryk and Raudenbush (1992) describe these statistics as indicators of educational effectiveness and equity. In other words, where school districts produce higher scores, controlling for the nested character of the data, socio-economic and organization factors, we infer that such districts are more effective than districts with lower scores. Additionally, we infer greater equity in districts that have a less strong or negative relationship between SES and achievement. Obviously, such inferences must be taken cautiously since numerous statistical matters, including the validity of the data and specificity of the modeling, affect the outcome. Thus, our purpose is not to make inferences about specific districts but rather to make inferences about the system as a whole, which we think is appropriate and defensible.

Each district's distribution of performance scores is characterized by two statistics: the intercept and the slope of the regression line.

The overall mean performance of schools, controlling for SES, was 259 points (up slightly from the unconditioned ANOVA model) with a standard error of 2.2; these are highly significant results. The average SES-performance slope was 21.1 with a standard error of 1.7, which is highly significant.* This indicates that one average, school measures of SES, is significantly related to both average levels of achievement as well as the slope of that relationship. The correlation (Tau as correlations) between slope and intercept is 0.69. In other words, there is a strong relationship between school-level measures of SES and performance.

The variance explained by the SCHLSES regression model is 25.7 percent. This result stands in sharp contrast to findings using OLS, which assigned more than 40 percent of the variance to school measures of SES.

According to Bryk and Raudenbush (1992), the problem with OLS methods is that they fail to properly account for the error terms associated with nested data. Thus, the problem is not only in the assignment of variance explained in the model but also in the specification of the regression coefficients. This record point is clearly illustrated in table 4, in which the regression coefficients for the affected school-level SES on performance are calculated using OLS and HLM. By comparison, the coefficients produced by HLM were generally about 20 to 25 percent larger than those generated by OLS, with the exception of the 1992–93 school year.

Our intention in reporting these findings is not to claim that OLS is inappropriate in all regression analyses; where there is relatively little dependence between levels within a data structure, OLS is appropriate and will produce

* The slope of the SCHLSES-performance relationship is relatively constant over time, suggesting that the findings reported for this year are not significantly different from other years, see table below:

	HLM random model	
	Coefficient	Standard error
1990–91	21.13	1.71
1991–92	20.12	1.70
1992–93	21.08	2.70
1993–94	20.89	1.95
1994–95	24.46	1.14

Table 4.—Comparison of Ordinary Least Squares (OLS) and Hierarchial Linear Regression Model (HLM) regression coefficients for school socio-economic status (SES) and school achievement

	Comparison of regression coefficients for school level SES			
	AVG OLS	HLM	Difference	As percent
SES90	17.45	21.13	3.68	21.1%
SES91	15.54	20.12	4.59	29.5%
SES92	22.15	21.08	-1.07	-4.8%
SES93	16.68	20.89	4.21	25.3%
SES94	21.80	24.46	2.66	12.2%

NOTE: Schools *N*=502; districts *N*=27.

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

results similar to those of HLM. Rather, our intention is to report the differences related to methods and the significance of recognizing dependence in nested data.

Additionally, we note that considerably less of the variance in performance data is explained by school-level SES measures using HLM than in predicted in OLS regression. This is good news for policymakers because it suggests that the bond between SES and achievement may be more malleable than has typically been assumed.

Now that we have described the variability of slopes between school districts, we seek to explain those variations relative to revenues and specifically by comparing the calculation of revenues by pupil count (REV/PUP) and weighted pupil count (REV/WPU). In other words, does the allocation of resources intended to equalize educational opportunities have an effect when educational outcomes are used as the dependent variable?

An Intermediate Two-Level Model: The Significance of Money Disappears

We have labeled this section an intermediate analysis because we only include one variable in each of the level 1 and level 2 regressions. The results are surprising because the interaction of school SES and district variables weakens the overall significance of the relationships, compared with the means as outcome, or random coefficient models discussed above. The results highlight an important message emphasized by Kreft and de Leeuw (Kreft and de Leeuw 1998), which is that the power of multilevel models is both its blessing and its bane. The results vary according to how one specifies the model. The only guide one can rely upon is theory. Fishing for results is not grounds for model building.

In the model displayed in figure 5, a level 1 predictor is introduced into the regression, while district SES is used as a level 2 conditioning variable for both the intercept and the slope intercept (the average SES-performance slope across all districts). SCHLSES was entered as raw data.

Table 5 displays the coefficient results for these regression analyses. In the first panel we see that when controlling for SES of schools (level 1) the influence of district SES on the intercept (average levels of achievement) is negative but significant. In other words, when district SES increases, the average level of achievement declines.* This is counter-intuitive and runs against the earlier analysis of district SES in the means as outcome HLM analysis. In contrast to the means as outcomes analyses, this two-level model provides evidence of the average direction of the slopes among the 27 districts included in this study. On average, as district SES increases, the relationship between school SES and performance increases positively and statistically approaches significance. Thus, high SES schools perform at a higher level in high-SES districts than in low-SES districts.

* The results of the regression that centered school SES differed significantly from the regression model that used uncentered data. Specifically the direction of the coefficient for the intercepts was positive instead of negative. Frankly, such a result makes more sense than the uncentered coefficient, which is negative. Currently, it is unclear what the uncentered model means. Kreft and de Leeuw acknowledge this problem explicitly and suggest that without strong theoretical reasons for centering, it should be avoided. Perhaps this advice is overly cautious but the general proposition of relying on a theoretical, rather than statistical premise for one's decision seems sensible.

Figure 5.—Intermediate intercepts and slopes Hierarchial Linear Regression Model (HLM)

Level 1 Model

$Y = B0 + B1 (SCHLSES) + R$

Level 2 Model

$B0 = G00 + G01 (DISTSES) + U0$

$B1 = G10 + G11 (DISTSES) + U1$

And then subsequently entering the following variables in seperate level 2 regressions using the above model.

(REV/PUP)

(REV/WPU)

SOURCE: Author's regression model.

These findings are disturbing because they indicate that the relationship between SES and achievement at the school level is augmented by the wealth of the district in which schools reside, an outcome contrary to the equalization goals of state school finance plans.

The second panel in table 5 displays even more disturbing results because they suggest that there is no statistical significance between the pattern of funding and either the level of school performance or the SES-performance slope. The pattern of results is not much better for the analysis of REV/WPU because the influence of revenues distributed in this manner is not significantly related to achievement. The only hopeful sign, from a policy perspective, is that the coefficient for this analysis is positive, suggesting the increases in funding by WPU are associated with increases in school-level performance.

Fourth Analysis: Full HLM Regression Model: The Significance of Money Reappears

In the above sections we have learned that in district SES revenues are significantly related to school performance. Additionally, we found that school SES is significantly related to performance. Finally, we discovered that when we add these variables into a simple two-level model, the significance of money all but disappears. District SES, however, has remained consistently significant in this analysis. In this section we explore the hypothesis that the interaction between school SES and district SES may confound the interpretation of the role revenues play in relation to school performance. The full model we examine is displayed in figure 6.

Thus, we believe that this more complex model should produce different results than the simpler model introduced above (see table 5). Specifically, we want to know whether district SES (DISTSES) and revenues per pupil (REV/PUP) significantly predict the intercepts, controlling for the SES of schools. In other words, once we control for difference in district SES, do revenues significantly predict performance? Also, do these variables significantly predict the within-district slopes?

Separate analyses compare the results for models including revenues per pupil and per weighted pupil unit separately. These results are presented in separate panels of

Table 5.—Intercepts- and slopes-as-outcomes for DISTSES, REV/PUP, REV/WPU, 1990–91

	Fixed effect	Coefficient	Standard error	T-ratio	P-value
District averages					
For INTRCPT1	DISTSES, G01	-12.951990	4.899802	-2.643	0.014
For SCHLSES slope	DISTSES, G11	4.560831	2.933783	1.555	0.132
Revenue/Pupil					
For INTRCPT1	REV/PUP, G01	0.000493	0.007365	0.067	0.948
For SCHLSES slope	REV/PUP, G11	-0.001045	0.005547	-0.188	0.852
Revenue/WPU					
For INTRCPT1	REV/WPU, G01	0.009489	0.019338	0.491	0.627
For SCHLSES slope	REV/WPU, G11	0.008882	0.013531	0.656	0.517

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

Figure 6.—Full intercepts and slopes Hierarchical Linear Regression Model

Level 1 Model
Y = B0 + B1 (SCHLSES) + R

Level 2 Model
B0 = G00 + G01 (DISTSES) + G02 (REV/PUP) + U0
B1 = G10 + G11 (DISTSES) + G12 (REV/PUP) + U1

And then subsequently substituing (REV/WPU) for (REV/PUP) in level 2 regression models.

SOURCE: Author's regression model.

table 6. The first panel describes the results for inclusion of revenues per pupil (SCHLSES is entered as an uncentered level 1 variable).

We are still interested in the general relationship between these variables and two outcomes: the average level of achievement (the intercept) and the average slope of regressions within the 27 districts. The relationship between district SES and the average level of achievement is negative and statistically significant. The effect of increased revenues and average levels of achievement is not significant; the direction of the coefficient is negative. These are not the effects policymakers anticipate.

The effect of these variables on the average slope of SCHLSES-performance relationship is more interesting. First, increases in the SES of the district is positively, and significantly, associated with a positive slope in the SCHLSES-performance relationship. In other words, there is a positive interaction between district and school measures of SES on performance that compounds the equalization goals of most state finance plans.

Controlling for district SES had an important and dramatic effect on the significance of revenues on school performance. In table 5, the explanatory power of revenues per pupil on performance measures was indistinguishable from the null hypothesis that there was no effect. When the variance of district SES is pulled from the model, direction of revenues changes and the effect of the variable approaches significance.

The same result pattern is evident for the measure of revenues per WPU, which suggests that the effect of vertical equalization, via the distribution of resources using a weighted allocation scheme, was masked by the effects of district SES.

Table 6.—Hierarchical Linear Regression Model statistics full model changed

	Fixed effect	Coefficient	Standard error	T-ratio	P-value
Revenue/Pupil					
For INTRCPT1	DISTSES, G01	-14.683383	5.128935	-2.863	0.009
For INTRCPT1	REV/PUP, G02	-0.005118	0.006913	-0.740	0.466
For SCHLSES slope	DISTSES, G11	8.999420	3.790910	2.374	0.026
For SCHLSES slope	REV/PUP, G12	0.009316	0.005939	1.569	0.130
Revenue/WPU					
For INTRCPT1	DISTSES, G01	-13.417124	4.911273	-2.732	0.012
For INTRCPT1	REV/WPU, G02	0.007296	0.017612	0.414	0.682
For SCHLSES slope	DISTSES, G11	7.681053	3.302069	2.326	0.029
For SCHLSES slope	REV/WPU, G12	0.020210	0.012622	1.601	0.122

SOURCE: Table calculated by author from data collected by the Utah State Office of Education.

These are important findings and illustrative of those predicted by recognition that school finance data are nested within a hierarchically structured organization. Failure to control for the interaction of level 1 and level 2 variables, failure to properly control for the error estimation within multilevel data structures, and failure to properly partition the within- and between-group variance in multilevel models raises serious questions about the statistical analyses of the relationship between educational resources and school effects

The last question we examine, the one that initiated this research, addresses whether the allocation scheme that attempts vertical equalization (REV/WPU) significantly improves educational outcomes compared with a per pupil allocation scheme (REV/PUP). We recognize that this is not the best test of this model but, for the moment, it is the best *post hoc* approach we can design.

The evidence suggests that there are not significant improvements, relative to school levels of performance, between the two models. Revenues distributed either by WPUs or by per pupil counts are not significantly different in their relationship to school performance measures. The deviance statistic, which is described by Bryk and Raudenbush as "a goodness of fit" statistic also confirm this general conclusion. Generally, the lower the deviance statistic the better the model fits the data, the less error there is in the estimation of population parameters. When one compares the deviance statistics for the full HLM regression (Intercepts and Slopes) used in these sections of the analysis, one can find a change of only two points. These figures are not calculated as representing a statistically significant different result (Chi Square calculation). Hence, we conclude from the available evidence that these two models are not significantly different: a finding running contrary to the traditional equity study we have mentioned as a source of motivation for this study.

[I]t was the exploration of these ideas using multilevel modeling that lead to the finding that, in this data, the relationship of expenditures per pupil to achievement was masked by the interaction of district wealth and school achievement.

Recalling the graphic illustration of the Federal Range Ratio, a traditional equity measure, the above results suggest that the difference in the trend line for the enrollment counts versus the WPU counts is not significantly related to either achievement levels or equity goals. Indeed, considering how these two measures mirror one another's ups and downs, it appears that the two measures are reflecting common influences rather than capturing different measures of activity. Thus, while the traditional equity analysis using measures of WPU suggest an improved equity picture, distributions accounted for by WPU are not substantially different than those by pupil count relative to performance.

Discussion and Conclusion

Multilevel models are appropriately applied to data that exhibits dependence between levels, data that is hierarchically nested. The statistical advantages of such regression techniques, compared with that produced by ordinary least squares, include better partitioning of variance between organizational levels (in this case the schools and their school districts), as well as better parameter estimates of intercepts and slopes. These are important advances in the statistical analysis applicable to the study of school finance and the relationship between educational resources and achievement data. Indeed, one of the important findings of this research was that as much as 35 percent of the variance in school-level achievement scores were accounted for by district variables (including differences in expenditure levels).

While the statistical advantages of multilevel regression models are important, what we found so intriguing was how thinking about the methods changed our thinking about the problem of school finance, equity, and productivity. We began to think, for example, in terms of cross-level interactions between district- and school-level characteristics. In some respects this point is obvious, which is to say that the performance of a school with a large number of high-needs students may differ depending upon whether it is situated in a resource poor and wealth school district. Nevertheless, it was the exploration of these ideas using multilevel modeling that lead to the finding that, in this data, the relationship of expenditures per pupil to achievement was masked by the interaction of district wealth and school achievement.

The point we take from this particular finding is that context matters. By context we mean to highlight the

BEST COPY AVAILABLE

organizational and policy environment in which educational services are produced and delivered to students. Ignoring the context in the study of the relationship between educational achievement and resources ignores exactly what is important to policy analysts. Quantifying monetary resources is, at best, a proxy for the potential of a school environment to produce quality services relevant to the educational needs of students. Context, in terms of organizational structure and policy, describe how people attempt to coordinate their activities around the use of available resources to some combination of purposes related to education. Factors related to context are the policy variables of interest, not the inert stockpile of resources in relation to a static measure of performance. The relationship of interest is the activity (organization and policy) between these two parameters (resources and achievement). Recognizing the hierarchical structure of school organization provided one means by which to explore these relationships. Applying multi-level regression techniques to these data refocused our attention and interest in ways we believe are useful and potentially powerful.

A second set of issues deserves concluding remarks. In this paper we began by recognizing the significant redistribution of state resources in an effort to promote school finance equity. We noted, however, that equity was a concern but not the goal of these efforts. Rather, legislators distributed money for equalization to promote and improve access to opportunities and consequently educational outcomes. The analysis discussed in this paper attempted to assess the productive contribution of school finance equalization among Utah's public schools. We compared the performance of schools with allocations based on simple pupil counts and the number of weighed pupil units (an account of funding equalization by comparable counts across school districts). Judgments about this comparison were based on a examination of intercept, slope, and deviance statistics. The intercept provides a comparison of achievement controlling for district and school characteristics. The slope provides a comparison of relationship between achievement and SES variables; an equity measure. The deviance statistic provides a "best fit" estimate of the two allocation models (expenditures per pupil and per weighed pupil).

The results found little evidence to support the hypothesis that the financing equity (the reallocation of resources to promote relative to need) was positively, or significantly, related to improved measures of achievement. This conclusion is conditional to the limits of the available data and their analysis; the results cannot be generalized beyond the Utah context. Nonetheless, these analyses provide an empirical referent by which to define the phenomena of interest and a set of methods by which to explore such questions.

While the potential of multilevel regression methods is great, it is important to recognize their limitations. Kreft and de Leeuw (1998), in their recent book entitled, *Introducing Multilevel Modeling* note that the power of multilevel modeling to more accurately capture and model reality is also one of its weaknesses. The problem is related to the robustness of findings. The sensitivity of multilevel regression to model specifications is great and, hence, slight changes in the regression model can have large changes in the findings. Kreft and de Leeuw argue that such a high-powered regression technique is not necessarily a good one for unguided exploration. Without a good theory guiding the analysis and a good understanding of the data used in the regression, it is very difficult to judge the merits of coefficients and *p*-values. The challenge of multilevel regression is testing the frontiers of modeling without losing sight of the theory guiding one's inquiry. The success of such a balance will require a statistically and methodologically informed researcher. Additionally, the utility of such research will require a statistically sophisticated audience.

The challenge of multilevel regression is testing the frontiers of modeling without losing sight of the theory guiding one's inquiry.

BEST COPY AVAILABLE

References

Berne, R. and Stiefel, L. 1984. *The Measurement of Equity in School Finance: Conceptual, Methodological, and Empirical Dimensions.* Baltimore, MD: The John Hopkins University Press.

Bridge, R. G., Judd, C. M., and Moock, P. R. 1979. *The Determinants of Educational Outcomes: The Impact of Families, Peers, Teachers, and Schools.* Cambridge, MA: Ballinger.

Bryk, A. S. and Raudenbush, S. W. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, CA: Sage.

Burstein, L. 1980. "The Role of Levels of Analysis in the Specification of Education Effects." In R. Dreeben and J. A. Thomas (Eds.), *The Analysis of Educational Productivity: Volume 1: Issues in Microanalysis.* Cambridge, MA: Ballinger.

Firestone, W., Goertz, M., and Natriello, G. 1998. *From Cashbox to Classroom: The Struggle for Fiscal Reform and Educational Change in New Jersey.* New York, NY: Teachers College.

Goldstein, H. 1995. *Multilevel Statistical Models.* London, England: Edward Arnold.

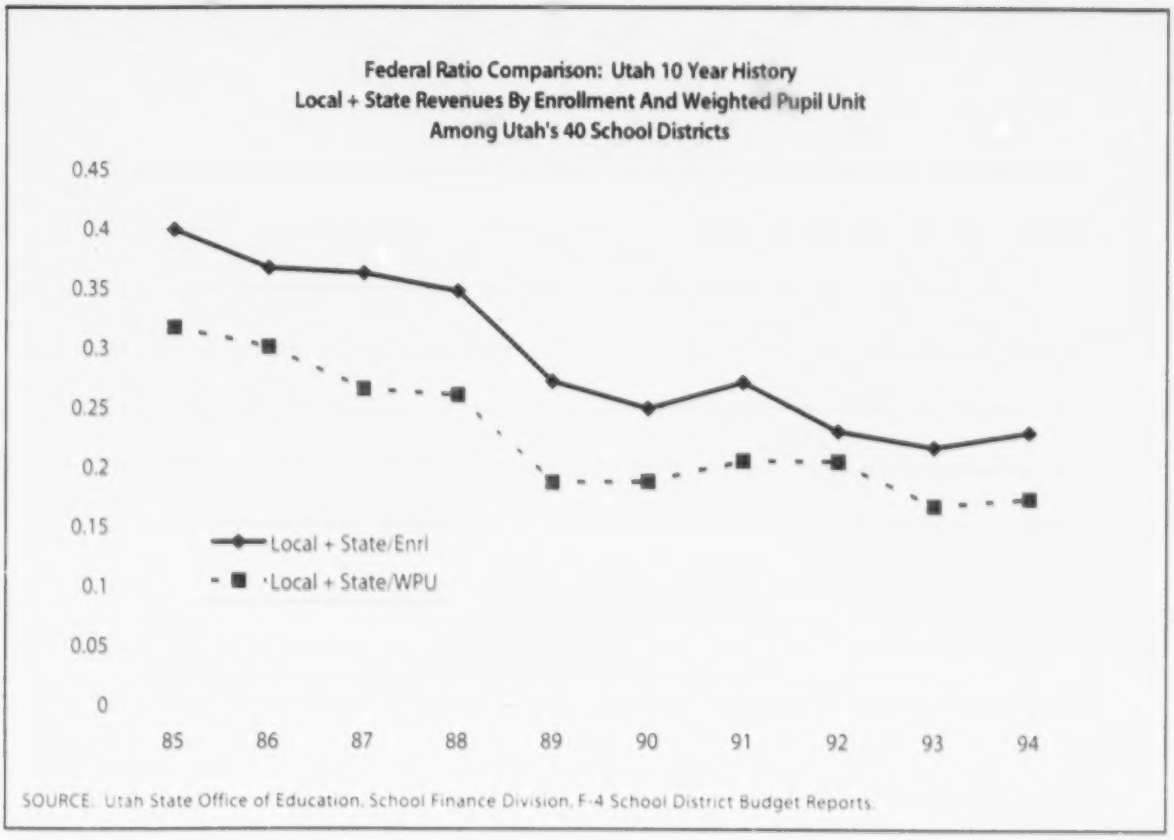
Kreft, I. and de Leeuw, J. 1998. *Introducing Multilevel Modeling.* London, England: Sage.

Monk, D. H. 1990. *Educational Finance.* New York, NY: McGraw-Hill.

Picus, L. O. 1995. *Does Money Matter in Education? A Policymaker's Guide.* In W. J. Fowler (Ed.), *Selected Papers in School Finance 1995.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Appendix

Graph of one traditional equity analysis for Utah



SOURCE: Utah State Office of Education, School Finance Division, F-4 School District Budget Reports.

Active Graphics Methods for the Analysis and Display of Education Data

Laurence Toenjes

University of Houston

About the Author

Laurence Toenjes is a research faculty member of the University of Houston's Department of Sociology. The research group with which he is associated is currently performing a review of patterns of promotions and retentions as related to student performance on the Texas Assessment of Academic Skills standardized test.

Dr. Toenjes has been an active participant in school finance analysis in Texas during the past decade, and has also participated, as a private consultant, in school finance policy analyses in several other states. He has created computerized school finance models for the states of Illinois, Texas, Missouri, and Nebraska.

A primary interest of Dr. Toenjes is the development and use of interactive computer graphics software to display and analyze school finance and student performance data and the use of such techniques to communicate findings to policymakers.

Dr. Toenjes received his doctorate in economics from Southern Illinois University.

BEST COPY AVAILABLE

Active Graphics Methods for the Analysis and Display of Education Data

Laurence Toenjes

University of Houston

Introduction

The purpose of this paper is to demonstrate how several computer-assisted graphical techniques, referred to as active graphics methods, can provide significant improvements in the manner in which education data can be explored and analyzed and research results communicated to policymakers and to the public at large. The discussion is in the following order: (1) a brief background on active graphics, (2) background and description of a particular software implementation of these methods, and (3) several examples of actual applications to the presentation and analysis of education data.

Remarkable improvements in personal computer hardware and software tools in recent years have made it possible to perform feats of data analysis only dreamed of in the not very distant past. These advances have occurred so rapidly that the effort needed to stay abreast of them seems to leave insufficient time to take advantage of what they offer. With respect to the needs of data analysis, interpretation, and display, the advances in and availability of high resolution color graphics on fast yet relatively inexpensive desktop and even lightweight portable computers are especially significant. These provide the physical means to effectively present the results of quantitative research to nonquantitative educators, policymakers, and even the public at large.

A consequence of the expanding availability and use of powerful microcomputers for administrative purposes has been the proliferation of data for research purposes. The National Center for Education Statistics (NCES), for example, produces a number of CD-ROMs each of which contain hundreds of megabytes of data on schools, school districts, staff, and students in the United States. A single CD-ROM can contain the equivalent of 300,000 printed pages of information. In addition to NCES, many state departments of education also make available, some over the Internet, massive amounts of information on their schools and school districts. In the state of Texas, for example, up to 11 gigabytes of student level test score data can be made available to researchers, with individual student identifiers replaced with pseudo-codes permitting inter-year matching of students.

Fortunately, the same technology that helps collect and disseminate these huge data sets can also assist in analyzing, sorting, and helping make sense of them. The current technology permits types of analyses that just ten years ago could only be performed by a lucky few who had access to the very expensive computers and peripherals then required. In addition to being able to carry out more sophisticated and extensive statistical analyses, current and affordable technology also offers more effective ways to communicate the results of such analyses. It is this possibility, as stated above, that is the principle interest of this paper.

Brief Background on Interactive Computer Graphics

Static graphic representations of data, in printed form, existed long before the advent of computers. Current graphics technology makes it easier to produce static graphs, to edit and refine them, and to reproduce them

BEST COPY AVAILABLE

on paper or transmit them electronically to be viewed or printed elsewhere. Some argue that this technology also permits a proliferation of mediocre and misleading material (Tufte 1983, 107). Such graphics still play an important role in conveying information and research findings. There are some excellent compilations of static graphics techniques with which data analysts should be familiar. Among these are Wang (1978), Tufte (1983, 1990), Cleveland (1985), and Jacoby (1997, 1998); some of whom also include discussion of some nonstatic graphics methods. By the late 1980s publications appeared that were entirely devoted to what are termed dynamic graphics and a related field referred to as statistical graphics. Significant works included Cleveland and McGill (1988), Buja and Tukey (1991), and Cook and Weisberg (1994). There are also numerous works in related areas referred to as exploratory data analysis (EDA) and data visualization techniques. The intent in this paper is merely to convey some appreciation of the acceleration of work in these areas and some of the sense of bewilderment and awe that one unavoidably feels when peeking into an exploding research community.

The concluding section of Jacoby (1998, 87) contains a brief but very informative introduction to what the author refers to as active display methods: "These constitute a different strategic approach to graphical data representations, rather than different types of displays per se. Active display methods rely on movement and real-time interactions between the analyst and the data (e.g., Becker et al. 1988; Young et al. 1993)."

The end result of most research, especially in the social sciences, is to inform and influence public policy.

The implications of the action-oriented aspect to the methods reviewed by Jacoby go well beyond the interactions between the analyst and the data, however. If the analyst uses the same interactive methods in presenting research findings to others as were used in the initial explorations and analyses, it is possible for the audience to experience the same sense of discovery and active involvement as that experienced by the researcher. This is a subtle yet potentially very significant aspect of active display methods. The end result of most research, especially in the social sciences, is to inform and influence public policy. Before a research finding can exert that influence, a policymaker or a policymaking body must become aware of the finding and be influenced by it in a positive manner. Properly utilized, active display methods, as described and illustrated below, can significantly contribute to the likelihood that policy research will receive the full attention of the policymakers for whom it is prepared.

This is perhaps an appropriate place to confront the problem, which this paper faces, of using ordinary language and static-graphic exhibits to convincingly describe the advantages of active-graphic methods. In commenting upon a paper contained in *Dynamic Graphics for Data Analysis* by Becker, Cleveland, and Wilks (1988, 60), Howard Wainer stated that "They have done a good job of conveying the general idea and many of the details, but the magic is missing." (Wainer 1988). Indeed, to miss the magic of these techniques is to miss the point. For it is the magic that can draw attention and have an impact upon both analysts as well as policymakers.

The solution to this problem of the inadequacy of the printed medium to convey the magic of active display methods, however, also has a technological solution. The appendix to this paper contains addresses, telephone numbers, and Internet addresses for contacting the vendors of the software discussed in Jacoby (1998, 90–96). In some instances, trial versions of the software can be downloaded. In addition, instructions are provided in the appendix for downloading a trial version of the software used to generate the displays discussed in this paper, with sufficient instructions and data to reproduce some of the displays and to interact with them.

In his overview of active display methods, Jacoby presents a table showing which of 39 graphical procedures are included in a dozen different software packages. He lists three techniques that perhaps most clearly set these methods apart from others: linking, brushing, and spinning. Just five of the twelve software packages he reviewed contained all three of these special features, also referred to as linked plot windows, plot brushing, and 3-dimension real-time rotation (spinning). Each of these three

BEST COPY AVAILABLE

procedures is described briefly, from the perspective of their potential usefulness in the analysis of education data.

Linking. Linking refers to methods by which multiple graphs are simultaneously displayed with the individual data elements in all displays connected in such a way that a change in an element in one graph will result in a change to the corresponding element in all the other graphs being displayed.

The simplest case of linked displays might be one in which two distinct scatterplots are created, with each pair of *x-y* data taken from the same observation or data record. This permits showing two different relationships among two pairs of attributes of the data. Identifying or highlighting an individual point in one of the scatterplots simultaneously identifies the location of the corresponding point in the other scatterplot; Stuetzle called these "connected scatterplots" (Stuetzle 1987, 472). The usefulness of linking will be illustrated below.

Brushing. Brushing refers to the method of controlling the position of an outline shape (the brush) on the monitor screen, such that the points falling within the brush's perimeter are visibly modified by a change in color, intensity, or shape. But most importantly, with multiple, linked views presented on the monitor, the linked points are simultaneously modified in corresponding fashion in all views as the brush is moved within the reference or active view. Usually the size and shape of the brush outline can be changed, although it is most commonly restricted to being a rectangle (see Becker and Cleveland 1987). A modified version of this technique will also be illustrated below.

[S]pinning ... attempts to break the restrictive two-dimension barrier of computer screens.

Spinning. Spinning refers to the ability to rotate a three-dimensional figure, often a scatterplot, around any one, or all, of its three axes. This technique was one of the first of the active display methods to undergo experimentation (Fisherkeller et al. 1974, 128). The spinning of three-dimensional plots attempts to break the restrictive two-dimension barrier of computer screens. While this technique is promising and can be very fascinating to observe, its effectiveness depends critically upon the quality of its implementation, referring both to the hardware and software used. For this reason, and because it is not clear that this technique would consistently help clarify presentations to persons not familiar with three-dimensional graphs, it will not be discussed further. Spinning was not incorporated into the software that was used to create the graphics presented below.

In addition to linking, brushing, and spinning, one other method appears to have been incorporated into most implementations of these techniques—the scatterplot matrix. A scatterplot matrix is just that, a matrix of individual, bi-variate scatterplots. Generally, several variables are selected and then individual scatterplots containing all possible pairs of those variables are displayed in the form of a matrix. An example, generated by SPSS, is shown in figure 1. When brushing and linking are also available, highlighting points in one scatterplot will similarly highlight the related points in all of the others.

While the scatterplot matrix can often be used effectively, some of the advantages it provides to an analyst, such as compactness, minimum space devoted to titles and labels, and quick viewing of multiple relationships, may tend to overwhelm and confuse the casual viewer.

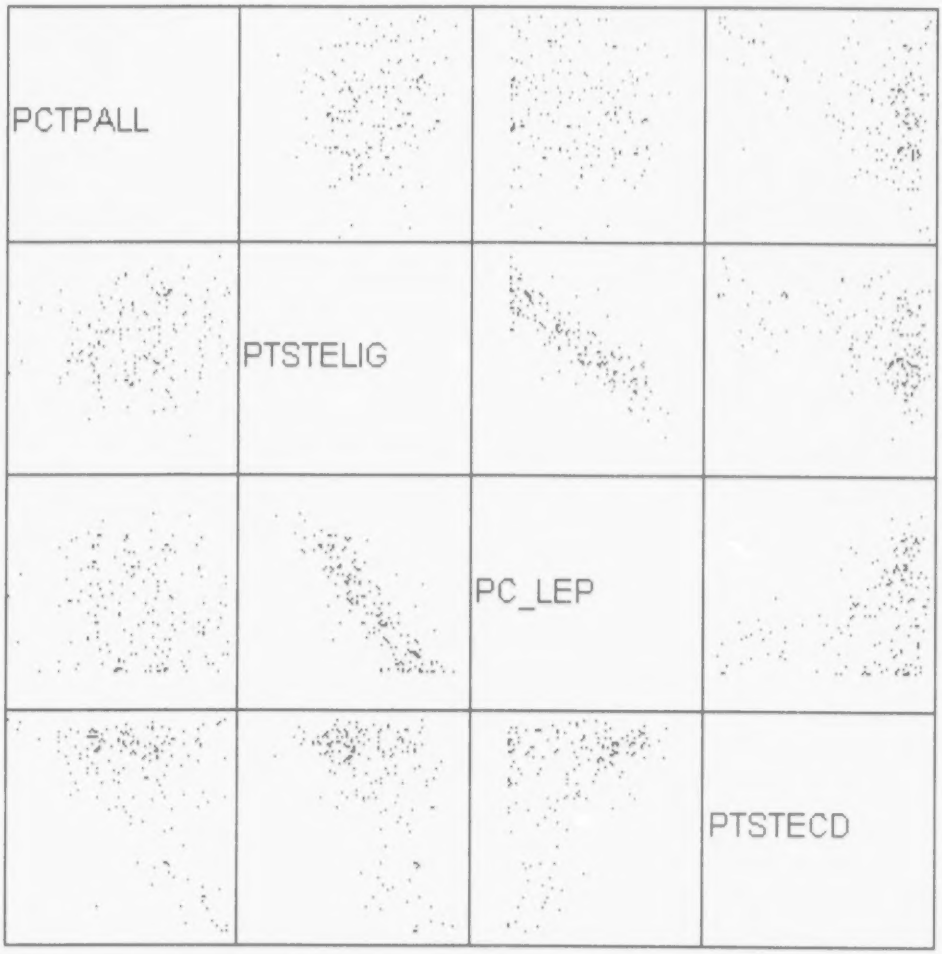
Implementation of Active Display Methods for use with Education Data

The specific software that will be used to generate sample displays shown and discussed below was, for the most part, developed independently of the progress described briefly above. A short history of the development of this software follows.

Being assigned to analyze school finance issues while working with the Illinois Bureau of the Budget during the years 1977–83, it became obvious to the author that one way to get the discussion beyond the prevailing district-by-district printouts of state aid was to use scatterplots depicting total revenue per pupil versus property wealth per pupil. Such diagrams provided a natural framework within which to discus school finance issues from the state's perspective (Toenjes 1982). But when using a scatterplot to make a point regarding some school finance issue or another, questions would inevitably arise

BEST COPY AVAILABLE

Figure 1.—Scatterplot matrix, Houston, Texas, elementary campuses 1997 Texas Assessment of Academic Skills (TAAS) Tests



NOTE: The variable names in the above chart have the following meanings:

PCTPALL: Percent of students passing all tests.

PTSTELIG: Percent of students enrolled in grades 3-8 taking the test and not exempted from the accountability system for reasons such as being classified as a special education student or of limited English ability.

PC_LEP: Percent of students classified as of limited English ability.

PTSTECD: Percent of students tested for accountability purposes who are classified as economically disadvantaged.

SOURCE: Texas Education Agency, data for Houston's 1996-97 elementary campuses.

regarding the locations of specific districts within the galaxy-like display (Illinois had more than a thousand public school districts at the time). Certainly it was necessary to know the location of the district associated with the City of Chicago, and some of the major downstate districts as well. At that time one felt successful if the scatterplot itself could be created, let alone identify individual points within it, on demand.

In 1989 the author worked with the Texas Center for Educational Research (TCER), with Dr. Catherine Clark and the late Dr. Billy D. Walker. We developed a school finance simulation model—the Texas School Impact

Model (TSIM)—that not only calculated state aid amounts due districts under various assumptions, but also displayed the results in the revenue per pupil versus wealth per pupil scatterplot diagram used previously in Illinois. However, the intervening years had seen a significant increase in the capability of the personal computer, in terms of speed of calculations and the quality of color graphics. Following are descriptions of some of the graphics features which were included in TSIM.

1. Individual points (school districts) in the scatterplot could be identified by capturing them within a box created by positioning a pair of crosshairs with the mouse, marking opposite corners of the box.

2. Upon identification, the points blinked in a contrasting color, and certain pre-designated data pertaining to the districts thus identified were written beneath the diagram.

3. With the point still blinking, a click of the right mouse button would permit the district's name to be written within the graph in a location selected by moving the mouse, with a narrow line connecting the name and the plotted point.

4. Again, with the point still blinking, and thereby focusing attention on that district within the diagram, pressing a particular key on the keyboard would cause a narrow stacked, color-coded bar to drop from the blinking point. The key to the color scheme used in the bar would also appear in the corner of the diagram. The relative lengths of the different colored sections of the bar were proportional to the percentages of the different types of real property in that district's tax base, i.e. residential property, oil-gas-mineral property, agricultural land and vacant lots, etc. This information was especially relevant at the time, due to the court suit that had successfully been brought against Texas' school finance system.[1]

[C]olored sections of the bar were proportional to the percentages of the different types of real property ...

5. The identification mode could be changed so that a district could be identified by entering the district's name. Features described in points 2–4 above were also available in the name identification mode.

6. A third mode, called the "Averages" mode, operated as follows: the mouse would be used to construct a rectangle around a number of points in the scatterplot. All of the points would then be highlighted in a distinctive color, the average values of the *y*-axis variable (total state and local revenue per pupil) and the *x*-axis variable (property wealth per pupil) calculated, and a small square with the sequence number of these calculations centered in it, drawn to the screen. The location of this identifying numeral was at the intersection of the average *y*- and *x*-axis values calculated for that subset of points. In addition, several numeric values—sums or weighted averages corresponding to certain characteristics of the districts just highlighted—were written to the data area beneath the diagram. These characteristics included total revenue per pupil, property wealth per pupil, total and operations tax rates, percentage of economically disadvantaged pupils, and the percentage of students passing the prior year's state standardized tests. With the rows numbered to correspond with the numerals placed on the screen within each subset of districts for which the calculations were performed, direct visual as well as quantitative comparisons could be made among different subsets of districts. For example, if a group of districts at each end of the wealth spectrum were highlighted in turn, immediate comparisons could be made of the basis of some of the crucial measures describing their respective financial and demographic characteristics.

7. A pattern of lines could be superimposed over the scatterplot of revenue versus wealth per pupil, in which each line corresponded to a specific tax rate. The points on each of these lines represented the locus of total revenue per pupil and wealth per

[1] The original finding against the state of Texas was in *Edgewood I.S.D. v. Kirby*, No. 362,516, 250th Dist. Ct., Travis Cty., Tex., 27 August 1987. This opinion was reversed in the Third Court of Appeals but upheld by the Texas Supreme Court in *Edgewood I.S.D. v. Kirby*, Tex. Sup. Ct., C-9353, 2 October 1989.

BEST COPY AVAILABLE

pupil combinations corresponding to that tax rate. With the respective tax rates displayed near these lines, it was possible to see at a glance the relationship between revenue per pupil, wealth per pupil, and tax rates. As different simulations were performed, the parameters to generate this family of tax rate reference lines would also be modified, so that the graphical presentation would remain current with the particular version of the school finance formula being used to generate the revenue data. A sample graph generated by this system is shown in figure 2.

8. Various methods of displaying only specified subsets of all districts were available. These included:

   (a) Districts within a specified enrollment range;

   (b) Districts within one or more selected counties;

   (c) Districts within a specified legislative district; and

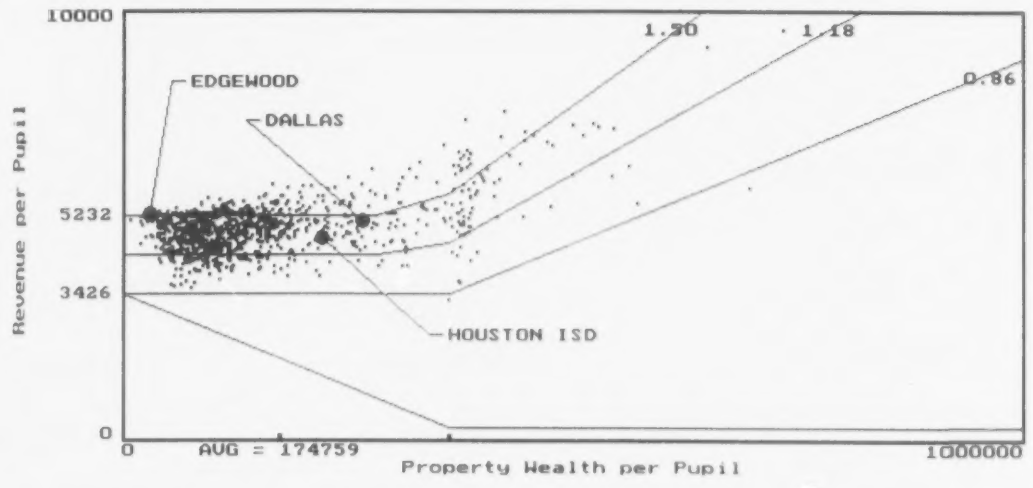
   (d) Districts which "gained" or "lost" under two different simulations.

9. Two different finance formula simulations could be run, and the increments in total revenue per pupil could be shown as colored vectors, with the origin or starting end for each positioned at the base value of revenue per pupil. These revenue change bars were colored green for positive increments, red for negative increments. Thus, the pattern of gainers and losers, their locations, and the amounts of the gains or losses per pupil, could be readily observed. In cases in which areas of interest were hopelessly overwritten, the vertical and horizontal scales could be reassigned, allowing the area of interest to be expanded. The mouse-controlled point identification technique (see item 1 above) was also available in this presentation, allowing instant identification of the district associated with any given change vector.

This school finance simulation and graphics display system proved highly effective and fairly popular. When installed on a portable computer, the conditions under which it could be used were greatly expanded, from meetings with individual legislators to rather large audiences, the latter utilizing various color projection systems. On one occasion it was set up in the Texas Senate chambers to augment a discussion of school finance issues presented by Drs. Walker and Clark, at the invitation of then Lieutenant Governor William Hobby.

The reason for recounting the above development and application of this particular system is to make the point that it was effective in large part because of the immediacy of its results, the relative simplicity of the graphics display, and the interactive manner in which policymakers'

Figure 2.—Texas school finance simulation, 1997–98



SOURCE: Texas Education Agency (data), simulations by author.

questions could be responded to, in most cases. Not only could relevant numeric data be quickly retrieved, but in addition, the graphical representation contributed significantly to an understanding of the issues being discussed.

TSIM was programmed using Borland's Turbo Pascal. In the late 1980s Borland incorporated objects, as in object-oriented programming, into their version of Pascal. A straightforward book on the topic by Alex Lane (Lane 1990) pointed the way to using a graph object to simultaneously produce multiple graphs using data from a common data set. This idea led to the realization that the graphical portion of TSIM might be contained in a graph object that could be replicated in numerous, different locations on the monitor, with each graph displaying different relationships among the data, yet all related (linked) in such a way that the points highlighted or blinked in one graph could be simultaneously identified in the related graphs. Thus, by combining the interactive features already incorporated in TSIM with the opportunities for generating and linking multiple graphs using object-oriented programming, a much more flexible graphics system became feasible.

[B]y combining the interactive features already incorporated in TSIM with the opportunities for generating and linking multiple graphs using object-oriented programming, a much more flexible graphics system became feasible.

The result of this process, first implemented under the DOS operating system, partially implemented on the Macintosh, and finally running under Windows 95, is a program called ScatterBrain. This program was used to generate the examples presented and discussed below. It is very likely that some of the other software packages described in Jacoby (1998) could be used in place of ScatterBrain. Several particular features of ScatterBrain will be mentioned, some of which are augmentations to the active display methods described in Jacoby (1998).

1. Linking is implemented, in that points highlighted or set to blink in one graph (the "active" graph) take on the same properties in all other graphs.

2. Brushing is not implemented as a fixed shape which can be moved freely across the screen, causing points to change colors as they fall within its boundaries. Rather, the mouse is used to establish opposite corners of a rectangle in the active graph. Any one of the maximum four displayed graphs can be designated as the active graph. When the rectangle is completed, the points inside it, and linked points in any other graphs being displayed, are all highlighted in the same color. Additional sets of points in the active graph can be selected, with each set highlighted in a different color. After seven such operations, the colors recycle.

3. Spinning is not implemented.

4. The scatterplot matrix method, as shown in figure 1 above, is also not implemented. In its place four different screen configurations are possible, showing one, two, three, or four linked scatterplots each. The graphs are individually defined and edited, with default or custom axes' labels and graph titles. Several types of reference lines can be overlaid on the graph including superimposed regression lines and associated lines representing plus and minus one or two standard errors of the estimate around the regression line. The specifications for each graph can be saved to a file. This permits using defined graphs in different combinations. Also, the screen locations of two graphs can be switched, by storing both and then recalling each to different screen locations. When all the graphs for a given screen configuration are edited as desired, the entire configuration, including the data display area (see item 5 below) can be saved into a "setup" file. This file records all of the major conditions and parameters that determine a given set of graphs and data table. Thus, a variety of screens can be configured, with different numbers of graphs, different variables plotted in each, different scales, or other options, and can be instantly recalled, used, and further modified, as desired.

5. An area below the graphs is reserved for the display of numeric data associated with the graph manipulations. This data display window is a central element in the ScatterBrain implementation and has been found to be essential when used with audiences of educators and policymakers. One of the unique aspects of education-related data, especially those pertaining to individual schools and

BEST COPY AVAILABLE

school districts, is that these entities have their own constituencies. Persons viewing discussions and displays of data pertaining to schools or school districts in which they are interested will usually have some relevant factual knowledge. If they see data presented that conforms to what they already know, confidence in the presentation is thereby instilled. Building on what is already known by at least some of those viewing the presentation, or often, in the case of policy members, by their staff, helps to move the discussion into areas with which they may not be so familiar.

The particular variables that are displayed in the data display window are selected from among all the variables in the data set in use. So, too, such properties as column width and number of decimal places are easily set. Such details can often contribute to, or detract from, the legibility of the display. When sets of points are highlighted, using the Averages mode, the user must specify beforehand which of the several aggregate functions are to be used to calculate the results to be written to the data table. Sums, simple averages, weighted averages, and a percentage calculation can be performed on any of the variables selected for display. While setting up the data table can be somewhat tedious, it adds enormously to the overall effectiveness of the combined graphics-data impact. When the data table selections have been determined, the specifications can be saved to a file so they are not lost. Modifications to the initial data table can be made and saved under an alternate name. Each can then be restored as appropriate. For example, different sets of variables could easily be displayed with the same graphs, for different purposes.

While setting up the data table can be somewhat tedious, it adds enormously to the over-all effectiveness of the combined graphics-data impact.

6. Geographical maps can also be incorporated into the displays. ScatterBrain considers a map as just another scatterplot with perhaps one or more special overlays. An overlay may represent one or more state boundaries, county boundaries, or for nongeographical data, a family of curves that help to clarify the relationships being discussed.

The U.S. Department of Education published the geographic midpoints (centroids) for most school districts in the United States (U.S. Department of Education 1990). These latitude and longitude coordinates can be used to plot school districts in any state or in all states, just as any other pairs of data elements are used to create a scatterplot. With ScatterBrain one of the four permissible scatterplots can be designated to be a map. This causes the program to look for one or more map files (up to five) which contain geographic boundary information to be drawn over (or under) the points in the scatterplot. Examples are presented below. Such maps are linked to any other graphs currently displayed, as described earlier. This makes it possible to combine the use of bivariate scatterplots with a geographic perspective provided by a map. Since many demographic and tax base variables have non-uniform geographic distributions, and since regional issues often play a major role in national, state, and local education issues, the possibility of explicitly providing a geographic perspective to these issues can be very informative.

7. Changes draw attention. For better or worse, changes to school finance formulas result in winners and losers. Or, if funded sufficiently, winners and bigger winners. In another arena, the increasingly common annual administration of statewide, standardized tests results in the need to assess which schools and school districts made adequate gains, and which, if any, fell behind. Both of these situations involve looking at incremental changes.

ScatterBrain includes a presentation graphic that highlights the pairwise change between two sets of points. In effect, the display superimposes two related scatterplots—test passage rates for two years plotted against the percentage of economically disadvantaged students, for example—and connects the pair of points for each district with a colored line. The line is green if the y-value increased, red if it decreased. The resulting diagram consisting of green and red change vectors quickly reveals any pattern to the incidences of gainers and losers. The point associated with the starting end of

BEST COPY AVAILABLE

each of the change vectors is linked to the corresponding points in the other graphs, map, or both. If a geographical map is in use, the regional dimension of gainers and losers can also be investigated.

8. Percent function. School finance equity debates often center on comparisons between property wealthy school districts and property poor school districts. Specifically, comparisons of total revenue per pupil for groups of school districts at the extremes of wealth per pupil, in which each group contains 5 percent or 10 percent of total state enrollment, are common. During the litigation in Texas, for example, there was a great deal of discussion as to whether the revised finance formula should be required to include those wealthiest districts containing 5 percent of statewide enrollment. (The Texas Supreme Court ruled that even the last 5 percent must be, for the most part, brought under the equalization formula.)

   A feature is built into ScatterBrain that permits forming the rectangles used to highlight and perform calculations on subsets of districts in such a way that it can be determined in advance what percentage of the total of a designated variable is contained in the rectangles before it is "closed." It is thus a simple matter to select a set of districts that include approximately any desired percentage of the state total number of students. Examples using this feature will also be demonstrated below.

School finance equity debates often center on comparisons between property wealthy school districts and property poor school districts.

Sample Applications

United States school districts: 1994–95 data

Points representing 13,742 regular public school districts are shown in figure 3. Aside from not including Alaska and Hawaii, the largest known omission consists of some 220 California districts. Geographical coordinates seem to have been omitted from the data set from which this information was drawn (U.S. Department of Education 1990). These 220 districts represented less than 10 percent of California's public school enrollments, however.

The non-geographic data used in this analysis were obtained directly from the National Education Data Resource Center, from data that are included on the 1994–95 Common Core of Data CD-ROM, published by the U.S. Department of Education, National Center for Education Statistics. The districts actually plotted in figure 3 are composed of 42,555,025 students. Approximately half of the 42.6 million students were enrolled in the 775 school districts whose total enrollments were 9,550 students or more. At the other extreme, only 6.26 percent of these 42.6 million students were enrolled in the 6,687 school districts that contained fewer than 1,000 students. The largest 100 districts which were mapped, those with enrollments of 40,750 or more, contained 9,509,251 students, just over 22.5 percent of the total (Anchorage, Alaska has 47,655 students, and Hawaii, a single district, would add 183,869 students to the total, if included). These 100 districts are plotted by themselves in figure 4. These large districts are very unevenly distributed throughout the country. Many states have none, while California has 10, Texas has 15, and Florida has 14, which total 39 in these three states alone. On the other hand, Illinois has only 1 (Chicago), while New York has just 2, New York City, and Buffalo.

An example of linked views involving a map and a scatterplot are shown in figure 5. Only districts with more than 500 students are plotted. Referring to the right-hand graph, the vertical axis represents the percentages of heads of household who have not finished high school, while the horizontal axis represents the percentages of heads of household who have at least one college degree. The shaded areas each represent districts containing approximately 5 percent of the U.S. public school enrollment, selected at opposite ends of the educational spectrum. The resulting pattern of the two sets of points, when highlighted on the U.S. map, generally show higher educational levels in the northeastern and upper mid-west regions, lower levels in the southeast, with interspersed enclaves in Texas and along the West coast.

The 4,367 districts with enrollments under 500 students comprised just 2.3 percent of total U.S. enrollments.

BEST COPY AVAILABLE

Usually the variation in the smaller districts is much larger than in the larger districts. Also, because their large number represents a relatively small proportion of the total student population, they can frequently be set aside with the result that the underlying relationship of interest is more discernable while still representing the vast majority of students.

The weighted averages for these two groups of districts, written beneath the graphs, are quite informative. For example, the districts with the higher proportions of heads of households with college degrees spent 44 percent more per student, had one-ninth the rate of childhood poverty, and had median incomes 2.9 times as great as in the districts with less educated heads of households. If the federal revenues per pupil are backed out of total expenditures per pupil, the more affluent set of districts would have spent 61 percent more than the others. Admittedly, these spending figures are not corrected for cost differences that might mitigate some of the difference in expenditures. But, the advantages to students coming from households with three times the income and seven times the rate of college degrees as the students in the opposite set only add to whatever extra financial resources their schools provide. This degree of inequity would not fare well if it were within the jurisdiction of a number of state courts which have ruled in favor of plaintiffs in several school finance suits this past decade.

Three State-Level Analyses: Indiana, New York, and New Jersey

Two views of Indiana school districts' data are shown in figure 6. The graph on the left shows expenditures per pupil on the vertical axis and the percentage of state and local funds from local sources on the horizontal axis. In the graph on the right are plotted the same measures of local household educational levels as shown earlier in figure 5 for the U.S. Districts containing approximately 13 percent of the students at each end of the measure of dependency on local funds are highlighted. The 29 districts most dependent upon local dollars, and hence most wealthy in terms of local resources, also have rates of college educated householders triple the rate among the 60 districts most dependent on state resources. Median household income in the more affluent districts is 39 percent greater than in the other set of districts and the incidence of childhood poverty is only approximately one-half as great. Expenditures per pupil are 29 percent greater, at $7,626 as compared to $5,897. All-in-all, the financial and social conditions among the two sets of districts are strikingly different.

Figure 3.—United States school districts

	NAME	FALLMEM	EXP PP	CHILDPOV	MED INC	MED HVAL	PC NOHS	PCCLGDG	FEDR PP	FEDR PCT
1	13742	42555025	6139	18	31051	93135	25	21	388	7

NOTE: The rationale for the distorted geographical depiction was to maximize the distances among the various points being plotted, thus minimizing the number of points that would be superimposed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1994–95 Common Core of Data

Figure 4.—Largest 100 public school districts in United States (excepting Alaska and Hawaii)

	NAME	FALLMEM	EXP PP	CHILDPOV	MED INC	MED HVAL	PC NOHS	PCCLGDG	FEDR PP	FEDR PCT
1	100	9509251	6333	23	30037	106132	25	25	531	8

NOTE: The rationale for the distorted geographical depiction was to maximize the distances among the various points being plotted, thus minimizing the number of points that would be superimposed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1994–95 Common Core of Data.

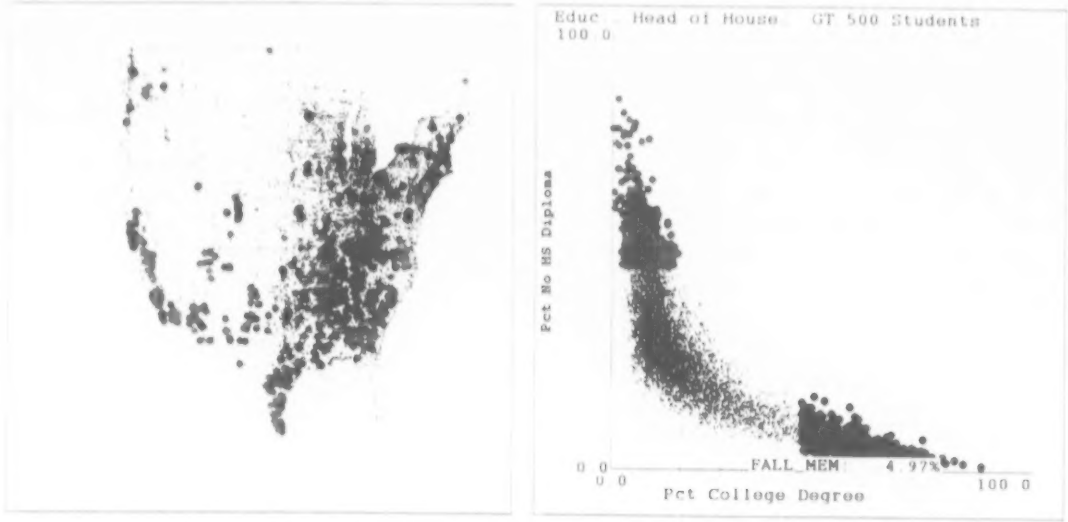
A comparable exhibit for New York State is shown in figure 7. In this case, the more affluent set on the right (group 1) spends 48 percent more per pupil, on average, than districts in group 2. Rates of childhood poverty differ by a 4:1 ratio, and median family income by more than a 2:1 ratio. Percentages of householders not having a high school diploma differ by 2:1, while the ratio of percentages with college degrees is nearly 3:1. All of these measures favor students in the more affluent group. In this case the 274,000 students in each group represent approximately 10 percent of New York State's total public school enrollments.

Two sets of districts, each containing approximately 111,000 students, were similarly selected in New Jersey. Again, each group represents approximately 10 percent of total statewide public school enrollment. This comparison is shown in figure 8. New Jersey differs from Indiana and New York in that the average expenditures per pupil are virtually identical among the two groups. State aid has effectively equalized the average expenditure levels between the districts at opposite ends of the local wealth spectrum. However, childhood poverty is 11times more prevalent in the less affluent group, median household income about one-third, and the percentage of heads of households with college degrees is one-fifth that in the more affluent group. Although average expenditure levels have been equalized, it is unlikely that equal revenues adequately compensate for the vastly different social and economic conditions in which the less advantaged students find themselves.

The scatterplot for New Jersey in figure 8 is repeated on the left side of figure 9. Two sets of districts are again highlighted in this figure, with each group in this case containing approximately 20 percent of New Jersey's total state enrollment. The same districts are also highlighted on the map of New Jersey on the right side of figure 8. In addition, several individual districts are identified.

The main point in presenting data for these three states—Indiana, New York, and New Jersey—is to demonstrate how quickly and dramatically issues can be meaningfully

Figure 5.—Educational levels of heads of household



	NAME	ST	FALLMEM	EXP PP	CHILDPOV	MED INC	MED HVAL	PC NOHS	PCCLGDG	FEDR PP	FEDR PCT
1	442	27	2098386	7919	4	54974	191727	7	54	157	2
	[illegible]	27	2079155	5479	36	18937	54557	53	8	666	12

NOTE: The rationale for the distorted geographical depiction was to maximize the distances among the various points being plotted, thus minimizing the number of points that would be superimposed.

SOURCE: U.S. Bureau of the Census, 1990 census data.

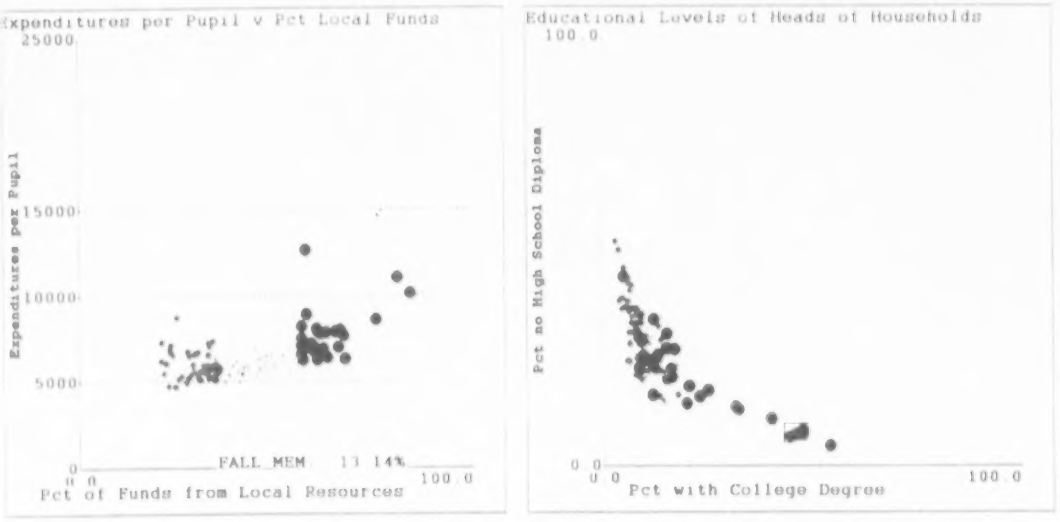
explored using active graphics methods. The ability to identify individual districts upon request can often make such presentations more compelling to an audience of stakeholders. The possibility of using data from numerous states to show similarities or differences among states can also be effective in making a point about a given state. The three states chosen are not unique. The data for most states will show interesting patterns. In fact, Hawaii, with a single, unified system, may be the only state for which statewide school district data, presented as above, would fail to reveal interesting contrasts and relationships.

"Substantially" equal dollars for equal effort in Texas

The principal focus of the series of court cases in Texas, beginning in 1987, was wealth neutrality. In particular, the degree to which all school districts had access to equal dollars for equal levels of tax rates was of paramount interest to the Texas Supreme Court. The graphs contained in figure 10 provide two views of the current system, as simulated with 1997–98 data. When completely phased in, no district is to have access to more than $280,000 of property wealth per weighted student (WADA). In the displays presented, WADA has been normalized back to average daily attendance (ADA), a smaller number.[2] In these terms, the maximum permissible accessible wealth would be closer to $370,000 per ADA. A vertical cluster of districts can be seen in the left graph of figure 10 near this value. Those to the right of this cluster have not yet been completely incorporated into this new regime through one of several acceptable methods of wealth divestiture. It should be noted that these districts, with higher wealth per pupil and generally higher total rev-

[2] Specifically, the WADA figures were multiplied by the ratio of the sum of WADA divided by the sum of average daily attendance (ADA), both summed across all school districts. This modified weighted attendance number is referred to as estimated daily attendance with the acronym EDA. Thus, in figure 10, TR_EDA refers to "Total Revenue per EDA," V_EDA refers to "Property Wealth per EDA." Other headings used in figure 10 are defined as follows: MTR—maintenance tax rate; TTR—total tax rate; ABA—adjusted basic allotment; AA—adjusted allotment; PDI—price differential index, a type of regional cost index. Some of these concepts only have relevance within the Texas school funding system.

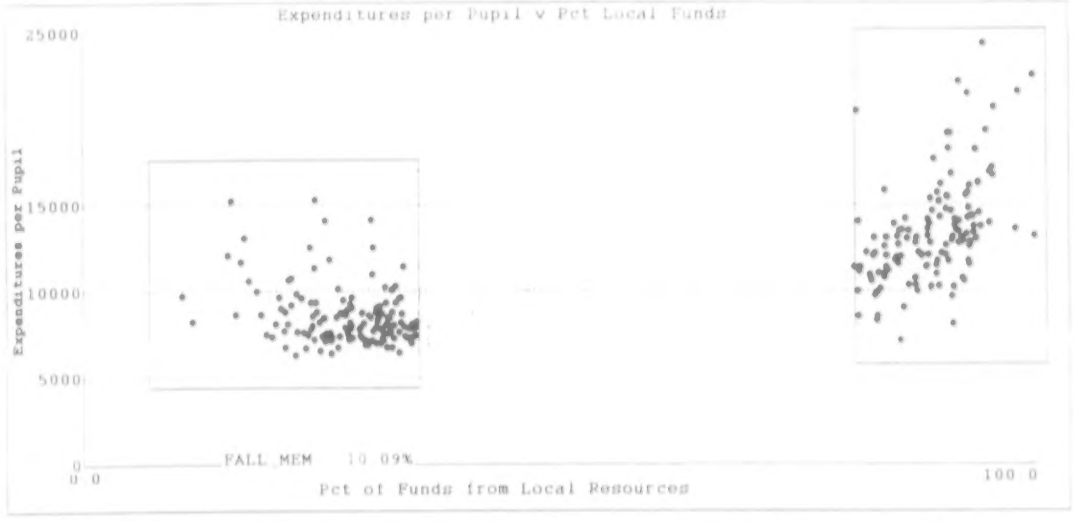
Figure 6.—Indiana school districts' expenditures per pupil and educational levels of heads of household



	NAME	ST	FALLMEM	EXP_PP	CHILDPOV	MED_INC	MED_HVAL	PC_NOHS	PCCLGDG	FEDR_PP	FEDR_PCT
1	29	18	126479	7626	10	35141	75303	17	28	279	4
2	93	18	127002	5897	19	25294	41786	32	9	316	5

SOURCE: U.S. Department of Education. National Center for Education Statistics, 1994–95 Common Core of Data.

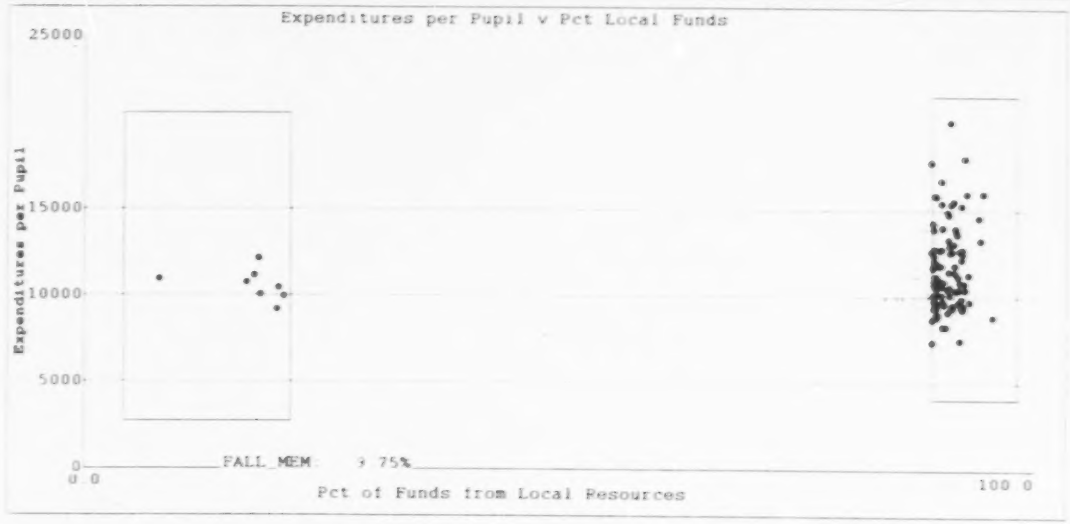
Figure 7.—Expenditures per pupil, state of New York public school districts



	NAME	ST	FALLMEM	EXP_PP	CHILDPOV	MED_INC	MED_HVAL	PC_NOHS	PCCLGDG	FEDR_PP	FEDR_PCT
1	134	36	274804	12688	5	57258	244875	14	41	176	1
2	151	36	274448	8559	21	25095	56014	28	14	480	5

SOURCE: U.S. Department of Education. National Center for Education Statistics, 1994–95 Common Core of Data.

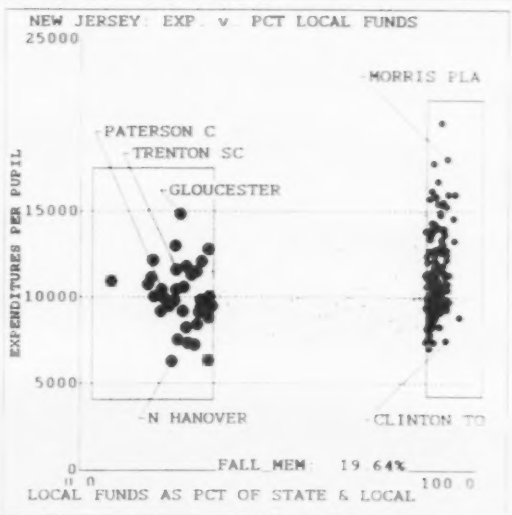
Figure 8.—Expenditures per pupil, state of New Jersey public school districts

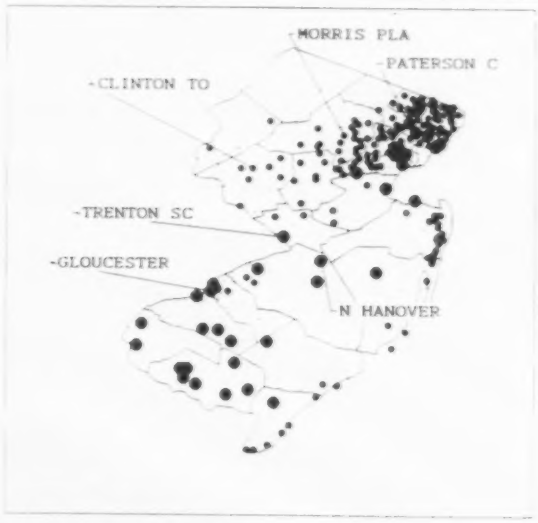


	NAME	ST	FALLMEM	EXP PP	CHILDPOV	MED INC	MED HVAL	PC NOHS	PCCLGDG	FEDR PP	FEDR PCT
1	101	34	111712	11158	3	62140	255718	12	46	125	1
2	8	34	111803	11235	34	23355	97360	46	9	748	7

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1994–95 Common Core of Data.

Figure 9.—New Jersey school districts, 1994–95



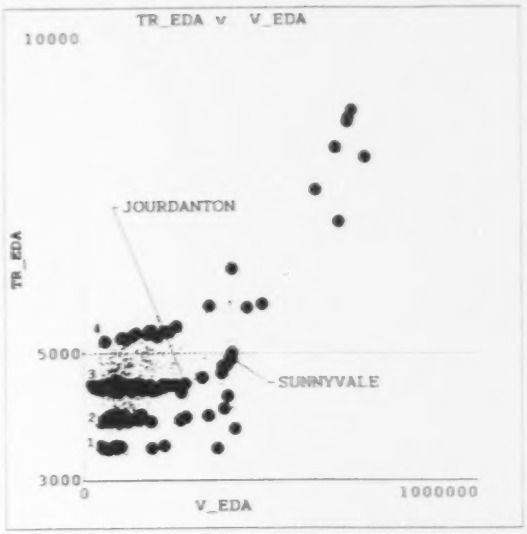


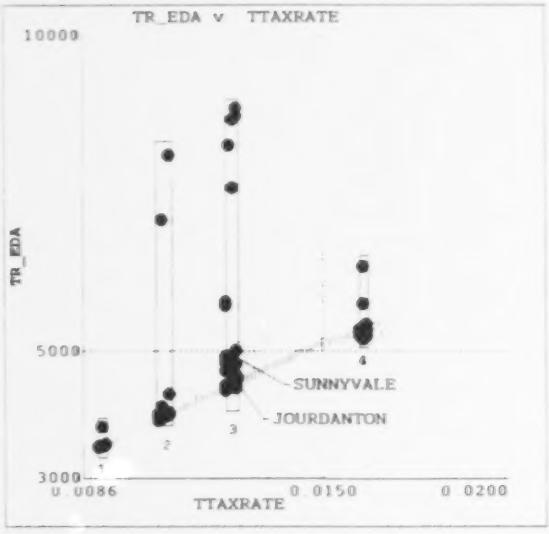
	NAME	ST	FALLMEM	EXP_PP	CHILDPOV	MED_INC	MED_HVAL	PC_NOHS	PCCLGDG	FEDR_PP	FEDR_PCT
1	16	34	111881	11170	3	64414	257647	12	44	125	1
2	142	34	225265	10702	4	59347	241116	13	45	143	1
TRENTON SCHOO		34	13634	11551	28	25719	70661	44	11	823	6
PATERSON CITY		34	22907	10379	28	26960	136311	47	9	821	3
MORRIS PLAINS		34	468	17983	4	64766	234548	11	47	165	1
CLINTON TOWN		34	459	7468	1	52851	191907	6	43	179	1
N HANOVER TWP		34	1746	8271	6	49422	166711	13	15	1578	20
GLOUCESTER CI		34	1849	14816	11	28998	62814	38	6	469	4

SOURCE: U.S. Department of Education, National Center for Education Statistics, 1994–95 Common Core of Data.

BEST COPY AVAILABLE

Figure 10.—Texas school finance formula, 1997–98 simulated data





	NAME	ADA	TR_EDA	MTR	TTR	EDA	V_EDA	ABA	AA	PDI
1	12	5654	3563	0.0091	0.0091	6807	155037	2508	3111	1.07
2	41	30706	4168	0.0100	0.0108	35654	169624	2499	2950	1.07
3	98	119447	4552	0.0114	0.0126	126044	166059	2554	2763	1.10
4	16	128761	5330	0.0129	0.0161	121910	194618	2628	2633	1.14
JOURDANTON		1120	4519	0.0120	0.0127	1270	258534	2506	2894	1.07
SUNNYVALE		334	4931	0.0126	0.0126	374	372884	2523	3327	1.08

SOURCE: Texas Education Agency (data), simulations by author.

enue levels, contain less than 2 percent of all students in the state.

The left-hand graph in figure 10 shows, for each district, total (state and local) revenues per pupil and property wealth per pupil. The graph on the right also plots total revenue per pupil on the vertical axis, but the horizontal axis represents total property tax rates. The four sets of districts were highlighted, in turn, by forming the rectangle to encompass districts in the right-hand graph with approximately the same total tax rates. The average tax rate for group 1 is 0.0091, for group 2 it averages 0.0108, for group 3 the average tax rate is 0.0126, and for group 4 the average is 0.0161. In the first three groups identified, most of the districts in each group plotted at approximately the same level of total revenue per pupil. For each of these groups, those districts that were noticeably above the majority in the right-hand view can be seen to be those with the highest levels of wealth per pupil in the left-hand view.

The district named Jourdanton, in group 3, is identified in both graphs. The horizontal location of this district in the left-hand view approximately marks the maximum wealth per pupil that is equalized in the second tier of Texas' equalization finance formula. Districts in group 3 to the right of Jourdanton define an upward-sloping line, terminating at the maximum wealth point, where Sunnyvale is identified. If desired, districts in this intermediate zone could be highlighted and the number of districts and the number of pupils in those districts tallied. The average difference in total revenue per weighted pupil guaranteed at opposite ends of this zone is estimated at $600 (assuming tax rates of 0.0150), and is one of the remaining areas of fiscal inequity in Texas' formula. As a result of recent changes to Texas' school finance formula, the major remaining reason for differences in revenue per pupil in Texas is due to variations in tax rates, as figure 10 helps to illustrate.[5]

[5] The author estimates that of the remaining variation in expenditures per weighted pupil in Texas, approximately four-fifths is attributable to tax rate variations, with about one-fifth due to differences in effective property wealth per pupil, after recapture is taken into account.

Another region that is not equalized in Texas' formula is defined by total tax rates exceeding 0.0150. The districts contained in group 4 are in this region, as each had a total tax rate well above 0.0150. Because the state only equalizes tax rates up to 0.0150 in the second tier, that portion of the rate above this level generates only local property tax dollars, at all wealth levels. Therefore, the line connecting these districts, as plotted in the left-hand graph, is decidely not horizontal. Equal dollars for equal tax rates does not hold true.

By interactively moving back and forth between the two linked graphs in figure 10 highlighting groups of districts or individual districts, it is possible to present a fairly accurate view of the relationships between property wealth, total tax rate, and total revenue per weighted pupil within Texas' school finance formula. The same can also be done for systems in other states. Central to these diagrams is the use of a weighted pupil count that approximately reflects the pupil weights, implicit or explicit, actually used in generating revenue entitlements within the finance formula.

Texas TAAS testing, testing, testing...

The Texas Assessment of Academic Skills (TAAS) has been in place in relatively unmodified form since the spring of 1994. Tests in reading and mathematics are administered in grades 3 through 8 and in grade 10 as part of the Exit test. Writing is tested in grades 4, 8, and also in the Exit test.

One of the major benefits of Texas testing program has been the ability to make comparisons between campuses and districts.

One of the major benefits of Texas' testing program has been the ability to make comparisons between campuses and districts. The hope is that programs and procedures associated with higher performances on some campuses, if verified, can be adopted by others.

The results on the 1998 TAAS are shown for 65 elementary campuses in Austin, Texas, in figure 11. The vertical axis represents the percentage of students on each campus who passed all tests. The horizontal axis represents the percent of test takers who were classified as economically disadvantaged. The majority of such students were so classified by their eligibility for the Federal free- or reduced-price lunch program.

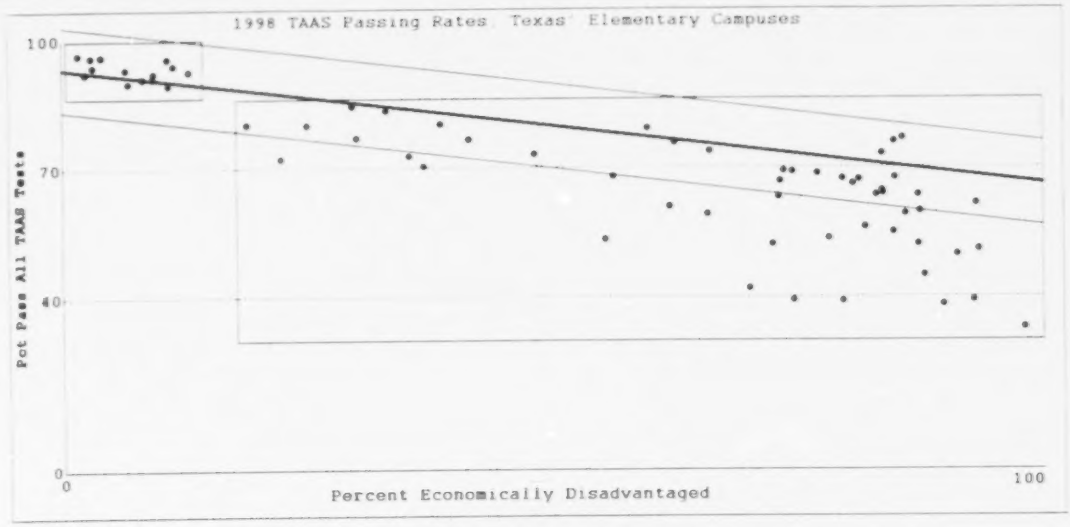
The thick, downward sloping line is the linear regression line fitted to this relationship for all elementary campuses in the state testing 50 or more students in 1998. The thinner downward-sloping lines are one standard error of the estimate (SEE) above and below the regression line. As can be seen, the cluster of 14 high socio-economic status (SES) campuses tend to be above the regression line, while the remaining 51 lower SES campuses definitely tend to be below the line. This is not an unusual pattern. A number of suburbs in the Houston area, for example, have similar patterns, with the high SES campuses tending to be within the SEE lines, but with their lower SES campuses tending to fall below the lower SES line. This raises the question as to which campuses, in which districts, are keeping the slope of the regression line up?

Aldine, whose elementary campuses are displayed in figure 12, is one such district. Keeping in mind that the regression line was fitted to all elementary campuses in the state testing 50 or more students for accountability purposes, all but 3 of Aldine's 32 elementary campuses have a positive residual, indicating higher passage rates than would be predicted by their levels of economic disadvantage.

The elementary campuses for both Austin and Aldine are highlighted, for comparison purposes, in figure 13. The larger circles represent Aldine's campuses, the smaller, Austin's. The two rows of data beneath the graphs help to compare these districts. Aldine has a higher test passage rate (84 percent versus 72 percent), more students who are economically disadvantaged (70 percent versus 51 percent), higher percentages of African American and Hispanic students (83 percent versus 63 percent, combined), and a greater proportion of limited English proficient (LEP) students (20 percent versus 13 percent). On the other hand, Austin tested 73 percent of its students for accountability purposes, whereas Aldine tested 68 percent. The higher proportion of LEP students in Aldine, many of whom are exempted in grades 3 through 5 if in bilingual education, could account for this difference in testing rates. The percentages of special education students are nearly

BEST COPY AVAILABLE

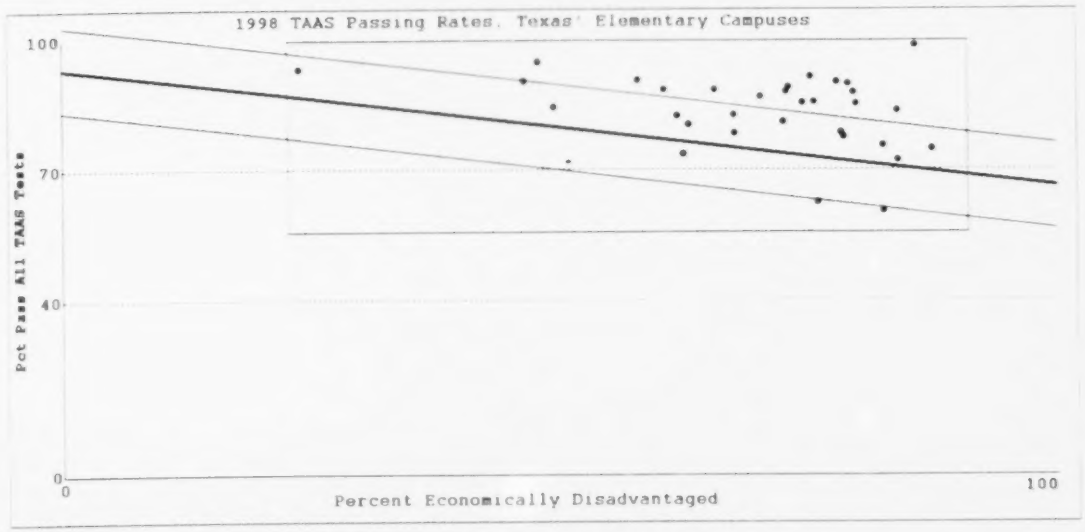
Figure 11.—Austin, Texas, elementary campuses' Texas Assessment of Academic Skills (TAAS) test results, 1998



	NAME	ENR310	PCTPALL	PCT ECDIS	PCT TSTD	BLACK	HISP
1	14	4864	93	7	85	4	13
2	51	13616	64	67	69	22	56

SOURCE: Texas Education Agency, 1997-98, Austin elementary campuses data.

Figure 12.—Aldine school district, Texas, elementary campuses' Texas Assessment of Academic Skills (TAAS) test results, 1998



	NAME	ENR310	PCTPALL	PCT ECDIS	PCT TSTD	BLACK	HISP
1	32	14456	84	70	68	36	47

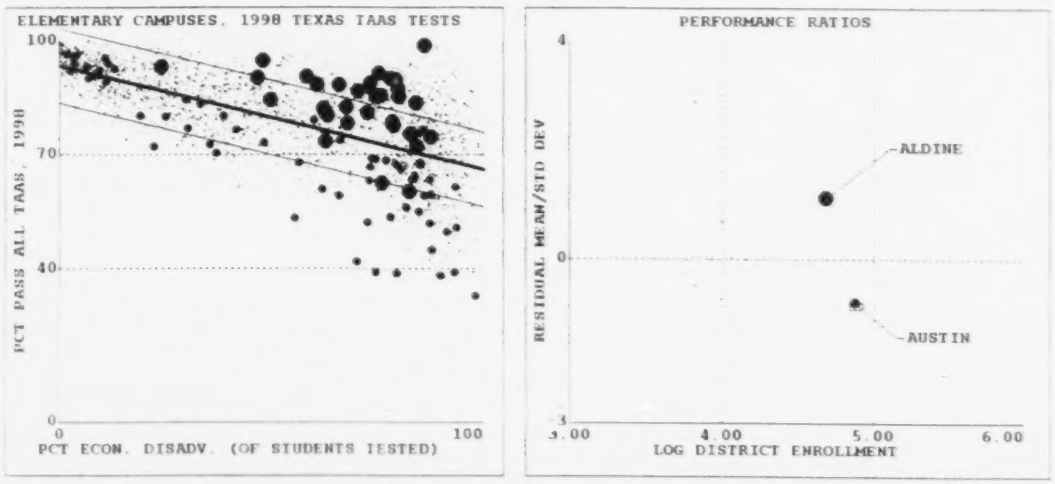
SOURCE: Texas Education Agency, 1997-98, Aldine elementary campuses data.

identical in the two districts, at 16 and 15 percent, respectively.

The variable plotted on the vertical axis of the right-hand graph of figure 13 requires explanation. For each district with 5 or more campuses testing 50 students or more for accountability purposes, an index was calculated, called the Performance Ratio (PR). The PR was calculated as follows. After regressing the percent passing all tests against the percentage of economically disadvantaged students, the mean and standard deviation of the residuals were calculated, for each district. The PR is the quotient of the mean of the residuals divided by the standard deviation of the residuals. For example, a district whose campuses were all well above the regression line, and by a fairly uniform amount, would have a relatively large and positive mean residual, and a relatively small standard deviation. Consequently, the PR for such a district, based on the performances of its elementary campuses, would be relatively high. On the other hand, a district whose elementary campuses' passage rates all fell below the lower standard error of the estimate line, would have a relatively large but negative PR. The idea behind this measure is that high or low indexes, at least for districts of comparable size or numbers of elementary campuses, might reflect some distinguishing district-level influence, positive or negative, as the case may be. The PR helps to identify such districts.

Once calculated, the PRs for each district were merged with the campus data records. In the right-hand graph of figure 13 these data, together with the logarithm of district enrollment, were used to plot the campuses. Since all campuses in each district had identical values of PR and the logarithm of district enrollment, all campuses for each district plot at the same point. Thus, when one of the points in the right-hand graph is highlighted, all of the corresponding campus points in the left-hand graph are simultaneously highlighted, revealing the pattern for all of the campuses in that district. Highlighting a second point in the right-hand graph reveals the pattern for another district's elementary campuses in the left-hand graph. This technique resulted in the two patterns seen in figure 13, contrasting Austin and Aldine in terms of the performances of their elementary campuses. The original graphs, as seen on the computer monitor, show the highlighted campuses for each district in contrasting colors, making the different patterns quite apparent. The right-hand graph, while somewhat interesting in its own right, might be called an instrumental graph, whose purpose is to help reveal a set of relationships in an associated graph.

Figure 13.—Austin and Aldine, elementary campuses' Texas Assessment of Academic Skills (TAAS) results compared



	NAME	ENR310	PCTPALL	PCT ECDIS	PCT TSTD	BLACK	HISP	PCSPED	LEP
1	32	14456	84	70	68	36	47	16	20
2	65	18480	72	51	73	18	45	15	13

SOURCE: Texas Education Agency, 1997-98, data for elementary campuses. Performance rates calculated by the author.

Four-Star Schools

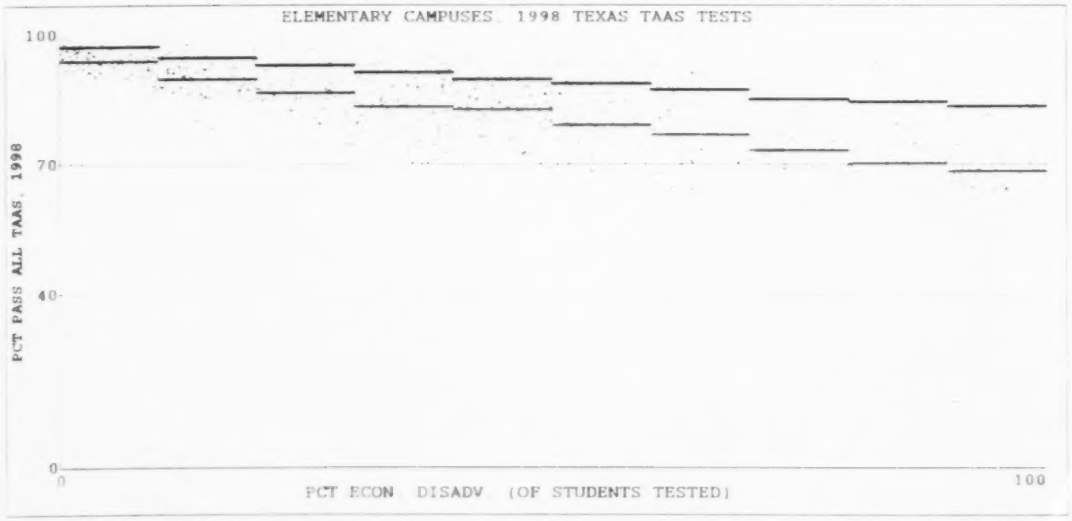
Texas is well known, some might say notorious, for its competitive high school athletic programs. It is likely that the designers of the public school accountability system anticipated that some of this competitive spirit might be directed towards academic achievement. To some extent, it appears that this has occurred. Marquees outside school buildings proudly display state ratings of "Exemplary" and "Recognized" status. The ratings are also finding their way into real estate information Internet sites. The Texas Education Agency publishes an annual report that rates all districts and campuses.

The Center for Houston's Future, a business-related, nonprofit organization, came to the conclusion that the rating system was overly complicated, and sought to develop one that was still meaningful but easier to explain. Dr. Darvin Winick and Dr. Toenjes were asked by the Center to try to design an alternative rating system for them. The results are represented in graphical form in figure 14. This diagram illustrates the performance of each elementary campus in the state which tested 50 or more students for accountability purposes, and permits assigning zero to four points or "Stars," on the following basis:

- One Star is awarded if the passing rate on all tests is at or above a minimum level—40 percent in figure 14.
- One additional Star is awarded if the passing rate on all tests is at or above 70 percent.
- One Star is awarded if, compared to all other campuses within the same 10 percentage-point range of economically disadvantaged students, the campus' passage rate is at the median for the group. The group medians are represented by the lower of the two stair-stepped lines.
- A fourth Star is awarded if, compared to all other campuses within the same ten percentage point range of economically disadvantaged students, the campus' passage rate is at or above the 85th percentile. The 85th percentiles are represented by the upper stair-stepped line.

Two Stars are thus awarded for meeting absolute levels of passage, regardless of the percentages of economically disadvantaged students. But, the other two take into account the proportion of economically disadvantaged students. It is evident from figure 14 that campuses with higher proportions of such students tend to have lower passage rates. Equally obvious is that there is great diver-

Figure 14.—Elementary campuses' Star rating criteria, 1998 Texas Assessment of Academic Skills (TAAS)



SOURCE: Texas Education Agency, 1997–98, data for elementary campuses.

sity in performance among campuses with comparable proportions of such students.

This diagram, as illustrated in figure 14, is used frequently. It is, of course, possible to show only the campuses in a particular district. When projected onto a large screen in front of an audience of principals and curriculum specialists for one of the major districts, it never fails to get their complete attention. It provides an effective framework to elicit discussion and raise questions as to why campuses fall where they do relative to one another. It also makes it possible to demonstrate instantly that some districts appear to overcome the negative factors that others use as excuses for not doing better.

Final Comments

This paper has been an attempt to describe with words and black and white static images some of the advantages of using interactive, high-resolution colored graphics methods to explore education data and to communicate findings to others. The methods described cannot replace the time-consuming and difficult statistical analyses and mathematical modeling that is always required to gain a better understanding of complex issues. But, active graphics methods can facilitate those tasks, and can almost certainly contribute to communicating results and insights to others in an exciting, compelling manner.

Appendix

The purpose of this appendix is to direct the reader to software that contains interactive graphics features such as those described above. The following table is taken from Jacoby (1998, 94). It includes names, addresses, telephone numbers and Internet addresses of eight vendors of such software.

Addresses and Contact Information for Software Vendors

Data Desk
Data Description, Inc.
803–573–5121
www.datadesk.com/datadesk/

JMMP, SAS/INSIGHT, and SAS/GRAPH
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
www.sas.com

S-Plus
Statistical Sciences Incorporated,
A Division of MathSoft
1700 Westlake Ave.
N. Seattle, WA 98109
www.mathsoft.com/splus.html

SPSS and SYSTAT
SPSS, Inc.
444 North Michigan Avenue
Chicago, IL 60611–3962
www.spss.com

Stata
Stata Corporation
702 University Drive East
College Station, TX 77840
www.stata.com

STATGRAPHICS Plus
Magnugistics, Inc.
2115 East Jefferson Street
Rockville, MD 20852–4999
ww.statgraphics.com

Statistica
StatSoft, Inc.
2315 East 13th Street
Tulsa, OK 74104
www.statsoft.com

ViSta
Forrest W. Young,
L. L. Thurstone
Psychometrics Laboratory
University of North Carolina
Chapel Hill, NC 27599–3270

Software and documentation can be downloaded from the World Wide Web at www.forrest.psych.unc.edu/ViSta.html

A trial version of the program which was used to create most of the graphs in this paper can be downloaded from the following World Wide Web site:

http://www.scatter-brain.com0

In addition to the software, a data set containing data for nearly 14,000 school districts in the United States can also be downloaded. Step-by-step instructions are provided to demonstrate how interactive graphs can be set up, experimented with, and printed out.

References

Becker, R. A., Cleveland, W. S. 1987. "Brushing Scatterplots." *Technometrics*, pp. 127–142.

Becker, R. A., Cleveland, W. S., and Wilks, A. R. 1988. "Dynamic graphics for data analysis." In W. S. Cleveland and M. E. McGill (Eds.), *Dynamic Graphics for Statistics*. Belmont, CA: Wadsworth and Brooks/Cole.

Buja, A., and Tukey, P. A. (Eds). 1991. *Computing and Graphics in Statistics*. New York: Springer-Verlag.

Cleveland, W. S. 1985. *The Elements of Graphing Data*. Monterey, CA: Wadsworth.

Cleveland, W. S., and McGill, M. E. (Eds). 1988. *Dynamic Graphics for Statistics*. Belmont, CA:Wadsworth and Brooks/Cole.

Cook, R. D., and Weisberg, S. 1994. *An Introduction to Regression Graphics*. New York: John Wiley and Sons.

Fisherkeller, M. A., Friedman, M. A., and Tukey, J. W. 1974. "PRIM9: An Interactive Multidimensional Data Display and Analysis System." In *Data: Its Use, Organization, and Management*. New York: Association for Computing Machinery, 140–145.

Jacoby, W. G. 1997. "Statistical Graphics for Univariate and Bivariate Data." *Sage University Papers Series on Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: Sage.

Jacoby, W. G. 1998. "Statistical Graphics for Visualizing Multivariate Data." *Sage University Papers Series on Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: Sage.

Lane, A. 1990. *Object-Oriented Turbo Pascal*. Redwood City, CA: M&T Publishing, Inc.

Stuetzle, W. 1987. "Plot Windows." *Journal of the American Statistical Association*, 82: 466–475.

Toenjes, L. A. 1982. "Illinois' School Funding Formula—A Mathematical and Geometrical Analysis." *Journal of Education Finance*, 8: 170–191.

Tufte, E. R. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Tufte, E. R. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press.

U.S. Department of Education, National Center for Education Statistics. 1990. *School District Data Book: 1990 Census School District Special Tabulation. U.S. Summary* [SDDB-01] (CD-ROM). Washington, DC.

Wang, P. C. C. (Ed). 1978. *Graphical Representation of Multivariate Data*. New York: Academic Press.

Wainer, H. 1988. "Comment: Deja View." in W.S. Cleveland and M. E. McGill (Eds.), *Dynamic Graphics for Statistics*. Belmont, CA: Wadsworth and Brooks/Cole, pp. 60–62.

Young, F. W., Faldowski, R. A., and McFarlane, M. M. 1993. "Multivariate Statistical Visualization." In C. R. Rao (Ed.), *Handbook of Statistics*, 9. New York: Elsevier.

Developing Student Resource Variables for the Early Childhood Longitudinal Survey

Lawrence O. Picus

University of Southern California

Lauri Peternick

American Institutes for Research

About the Authors

Lawrence O. Picus is a professor and Chair of the Department of Administration and Policy in the Rossier School of Education at the University of Southern California. He serves as the director of the Center for Research in Education Finance (CREF), a research center in the School of Education at the University of Southern California. CREF research focuses on issues of school finance and productivity.

Dr. Picus is past-president of the American Education Finance Association. Picus is the co-author of *School Finance: A Policy Perspective: 2nd Edition* (McGraw-Hill, 2000) with Allan Odden, and of *Principles of School Business Administration* (ASBO 1995) with R. Craig Wood, David Thompson and Don I. Tharpe. In addition, he is the senior editor of the 1995 yearbook of the American Education Finance Association, *Where Does the Money Go? Resource Allocation in Elementary and Secondary Schools* (Corwin 1995). He has also published numerous articles in professional journals.

In his role with CREF, Picus is involved with studies of how educational resources are allocated and used in schools across the United States. He has also conducted studies of the impact of incentives on school district performance. Picus maintains close contact with the superintendents and chief business officers of school districts throughout California and the nation, and is a member of a number of professional organizations dedicated to improving school district management. He is a member of the Editorial Advisory Committee of the Association of School Business Officials, International, and he has served as a consultant to the National Education Association, American Federation of Teachers, the National Center for Education Statistics, WestEd and the states of Washington, Vermont, Louisiana, Wyoming and Arkansas.

Prior to coming to USC, Picus spent four years at the RAND Corporation where he earned a Ph.D. in Public Policy Analysis. He holds a Masters Degree in Social Science from the University of Chicago, and a Bachelor's Degree in Economics from Reed College.

Lauri Peternick is a Senior Research Scientist at the American Institutes for Research. Her areas of specialization

BEST COPY AVAILABLE

include education finance and organizational behavior. Dr. Peternick earned her Ph.D. from the University of Chicago. She has a Bachelor's Degree in Political Science from Columbia University and a Master's Degree in Education Policy Analysis from Stanford University. In addition to her work on developing student resource variables for the Early Childhood Longitudinal Survey, she has presented, written, and co-authored papers on equity litigation, cost-of-education indexes, student-level resource cost models, site-based decision making, and school-level data collection. Recently, she has published two articles in the *Journal of Education Finance*: "Ohio: A Case Study in School-Level Data Collection," winter 1998, and "Site-Based Budgeting in Fort Worth, Texas," spring 1998.

Developing Student Resource Variables for the Early Childhood Longitudinal Survey

Lawrence O. Picus

University of Southern California

Lauri Peternick

American Institutes for Research

Introduction

Why Collect School-level Financial Data?

Spending on public K–12 education in the United States approaches $300 billion a year (U.S. Department of Education 1999). Despite this substantial sum of money, there is considerable doubt about how effectively our schools spend this money and whether or not additional resources will lead to improved student outcomes. Picus (1997) suggests that the real issue may be how the money is spent, not how much money is spent. One of the problems researchers have faced in the past is the difficulty of linking expenditures to individual schools and students. This has hampered our ability to understand how resources are linked to measures of student achievement.

Most fiscal data collected to date have been at the district and state level. At the time of publication, approximately eight states either collect or are in the process of developing systems to collect school-level fiscal data. These include Florida, Louisiana, Mississippi, Ohio, Oregon, South Carolina, Texas, and West Virginia. Florida, Texas, and Ohio have had the most experience in the collection of school-level data. However, Mississippi, South Carolina, West Virginia, and Oregon have also begun collecting school-level fiscal data. In contrast, Louisiana is in the process of establishing systems to collect school-level data. Washington recently considered this issue, but decided not to collect fiscal data at the school-level for the time being.

Developing school-level collections is important because district-level fiscal data may be masking the individual differences that exist across schools and among classrooms within schools. If how money is spent or how resources are used does not matter in terms of student achievement, then it is important to understand how resources are used at the school and, if possible, classroom, and student level. A new NCES study, the Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K), appeared to offer the perfect opportunity to begin collecting student-level resource data to help answer these complex and important questions.

The purpose of this paper is to describe the rationale and framework for collecting student-level resource data through the ECLS-K. At publication, the ECLS-K is in its second year of data collection. The vast majority of the study's kindergartners are now in first grade. However, prior to the first year of data collection for the ECLS-K, we proposed to NCES that the study include an additional set of measures for collecting student-level resource data. This article is a descriptive account of the item

development process we undertook several years ago and our rationale for including items. The paper also explains why the ECLS-K made (and still makes) an ideal vehicle for collecting student-level resource data.

We begin the paper with a brief description of the ECLS-K and its components. This is followed by a brief literature review describing both what is known about the allocation and use of educational resources and how studies of student resources can be used. Next, we present the conceptual framework that describes how this work fits within the ECLS-K and provides a model for collecting these cost data. Two final sections discuss the practical implications of adding additional student-resource items to the ECLS-K and our evaluation of the ECLS-K survey instruments. Again, our evaluation of the ECLS-K instruments and proposal to include additional student-resource items in the questionnaires occurred after the questionnaires had been developed and field tested, but before data collection for the main study had begun. We conclude the article by showing the actual NCES ECLS-K student-level and school-level resource instruments.

Overview of ECLS-K

The new ECLS-K offered an opportunity to collect new finance data that could help answer questions about the link between student outcomes and resource allocation and use. The ECLS-K is a comprehensive study of 23,000 kindergartners and their education. Beginning in the 1998–99 school year, the study will follow and measure kindergartens' academic environments, opportunities, and achievements through fifth grade.

The data collection design called for administering a set of assessments to members of the kindergarten cohort twice during the 1998–99 and 1999–2000 school years—in the fall and the spring—and once in the spring of every other year thereafter until fifth grade. Additionally, in the first year of the study, data were collected in both the fall and the spring of the 1998–99 school year from kindergartners' parents and teachers. Questionnaires were also administered once in the spring to students' school principals and special education teachers or service providers. Field data collectors also collected archival records data (e.g., information from students' records on absentees, tardies, grades). In subsequent years, questionnaire data are to be collected in the spring at two-year intervals from parents, teachers, and principals. Table 1 presents the data collection schedule for the ECLS-K.

Thus, the following survey instruments are to be administered several times during the course of the ECLS-K:

1. Teacher Questionnaires,
2. Student Information Checklist (completed by teachers,
3. School Principal Questionnaire,
4. Special Education Teacher or Service Provider Questionnaire, and
5. Parent Questionnaire.

With regard to education finance research, the goal was to gain an understanding of the resources invested in the child and their effect on the child's academic learning, achievement, and growth. The Teacher Questionnaire

Table 1.—Data collection schedule for the ECLS-K

	School year											
	Kindergarten		1st Grade		2nd Grade		3rd Grade		4th Grade		5th Grade	
Data source	Fall 1998	Spring 1999	Fall 1999	Spring 2000	Fall 2000	Spring 2001	Fall 2001	Spring 2002	Fall 2002	Spring 2003	Fall 2003	Spring 2004
Student												
Parent												
Regular teacher												
Special education teacher												
School principal												

* In the fall of the first grade, data will be collected from a 25 percent subsample of cohort members. For these children, assessments will be conducted and children's parents will be interviewed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K).

and the Special Education or Service Provider Questionnaire provide information as to the child's experience in the classroom. Information requested includes the nature of the instruction provided, books and other tools used during instruction, and the qualifications and experience of the teachers themselves. The Student Information Checklist provides insight into the background and abilities of the child as well as resources that have been and/or continue to be available to the child. This may include special programs or classes the child has taken or continues to take. The School Principal Questionnaire helps to provide a framework within which the child receives instruction. This includes insight into the resources of the school as a whole and more specifically what is available to children in the identified child's particular classroom. Lastly, the Parent Questionnaire provides insight into the home life of the child as well as the resources that the child received prior to beginning kindergarten. Insight into the home life includes, among other things, information about the child's reading habits and the existence of computers within the home.

Upon reviewing drafts of the ECLS survey instruments, it became apparent that by adding a few questions to each survey instrument and creating a new questionnaire to retrieve salary and benefit information from school business officers, researchers would have a new and rich source of data. This new data set could provide insight into some of the most pressing and complex issues in education finance today. However, before discussing the data elements that should be collected, and how they could be collected within the existing ECLS framework, it is helpful to review some of the current literature on the allocation and use of educational resources and how data on school- and student-level resource allocations might be collected and used.

The focus of most state finance reporting systems is on fiscal accountability, not understanding how or why resource decisions are made.

Literature Review

Despite the large sums of money spent annually for K–12 education, we know remarkably little about how those funds are used at the individual student and school levels. School finance studies have traditionally focused on school districts as the level of analysis, and most states only collect information from constituent school districts at the district level. The focus of most state finance reporting systems is on fiscal accountability, not understanding how or why resource decisions are made. These systems generally focus on object-level reporting. As a result, we know a great deal about how much our schools spend for salaries, benefits, contracts, etc., but relatively little about expenditures by function (instruction, administration, pupil services, maintenance and operations, transportation, etc.), and even less about how much is spent by individual programs (vocational programs, regular programs, and special programs for the mentally retarded, physically handicapped, emotionally disturbed and learning disabled.)

For example, many districts cannot tell us how much is spent per pupil for elementary versus secondary instruction, much less answer a question like what are per pupil costs for mathematics instruction at the secondary-school level, or how much is spent on individual students at the elementary level. Yet, until we can identify these costs, it seems unlikely we will be able to ascertain how the use of educational resources is linked to student achievement. The ECLS-K provides an excellent opportunity, because it begins with such a young cohort, to collect data over time on student resources and student outcomes.

Berne and Stiefel (1997) argue that student resource studies can answer three types of questions. They are:

- Resource effectiveness questions
- Equity questions
- Resource intent questions.

Each is described in more detail below.

Resource Effectiveness Questions

A large body of literature, both in economics and school finance, has focused on production function analyses that attempt to relate inputs to outputs. Studies of this type are useful for answering questions on the effectiveness of resource use, and the cost-effectiveness of different programs. To date, production function analyses that attempt to relate student outcomes to resources have not clearly identified a link between spending and student

BEST COPY AVAILABLE

achievement. Eric Hanushek's work in this field led him to conclude that there does not appear to be a systematic link between student achievement and level of spending (see for example, Hanushek 1989; 1994a; 1994b; 1996a; and 1996b). He does not suggest that such a link does not exist, only that at the present time, schools need to spend the resources they have more efficiently if they are to improve student learning with more money (see in particular, Hanushek 1994b).

In recent years, a number of authors have challenged Hanushek's findings, arguing that more money does relate to higher levels of student achievement. Hedges et al. (1994a and 1994b) have argued extensively that if different statistical methods are used to conduct meta-analyses of production function studies, there is a clear link between spending and student achievement (see Greenwald et al. 1996 and Laine et al. 1996). Ferguson (1991) found that "hiring teachers with stronger literacy skills, hiring more teachers (when students-per-teacher exceed 18), retaining experienced teachers, and attracting more teachers with advanced training are all measures that produce higher test scores in exchange for more money." Other work by Ladd and Ferguson (1996) in Alabama found similar links between spending and student achievement.

The ECLS-K is one of several NCES studies that collects student-level data from a sample of individual students and which, when aggregated, provide a reasonable representation of a particular grade cohort at the school level.

Cost-effectiveness studies are less common in the educational literature. In part this is due to the difficulty in measuring educational outcomes consistently across children. Cost-benefit analysis of which cost effectiveness is a derivative (see Levin 1983) relies on the ability to value both costs and benefits in dollar terms. The difficulty in education is that to compare student achievement, we need to rely on various test scores and measures of gain. Since tests in different subjects use different scales, as do different tests of the same subjects, it is virtually impossible to compare the cost effectiveness of different programs with district- and state-level aggregate cost data.

Berne and Stiefel (1997) argue that studies like the ones described above "...could be done with much more accuracy if there were student-level resource measures that were defined to be inclusive and to differentiate between kinds of programs and students. The data would be useful if it were gathered at the school level or, if it included a sample of individual student-level data that was representative at the school level." The ECLS-K is one of several NCES studies that collects student-level data from a sample of individual students and which, when aggregated, provide a reasonable representation of a particular grade cohort at the school level.

Equity Questions

School finance has a long history of analyzing funding equity. However, most of that work has looked at spending differences across school districts. Very few studies have considered school-level finance equity either within districts, or across districts in an individual state. Hertert (1996) analyzed school-level equity in California, but to do so was forced to collect data from a sample of school districts and enter their data by hand. Nakib (1996) analyzed school-level equity in Florida using that state's extensive school-level data. Picus (1993a; 1993b) used a national sample of school districts merged from the Schools and Staffing Survey and the 1987 U.S. Bureau of the Census, Census of Governments to analyze school-level expenditure patterns by various district characteristics such as size, location, and wealth. However, outside of these studies, there have been few school-level analyses of finance equity.

Berne and Stiefel (1997) suggest that "...a well-defined set of student resource variables would improve equity studies at the school level including studies that use administrative data, particularly if those variables are capable of serving as models for other data sets." Here again, the ECLS-K provides an excellent opportunity to begin collecting school-level data on per pupil spending.

Resource Intent Studies

The third category of questions Berne and Stiefel identify address how resources are used or how they flow to programs or schools. Resource Intent Studies are best conducted using the frameworks established by Chambers and Parrish's Resource Cost Model (RCM) and Cooper's InSight accounting model. Studies using these models provide a wealth of information on how educational resources are used. However, data collection meth-

BEST COPY AVAILABLE

ods are expensive, and all suffer from the inherent incompatibilities in the way districts and states report fiscal data. These complexities, combined with the need to make difficult decisions about allocation of overhead costs and central office expenditures have led most analysts to shy away from such efforts. The ECLS-K provides an opportunity to examine measures of program costs for a representative group of school children as they move from kindergarten to the fifth grade.

Resources Available To Children From Other Sources

An important component of resource availability for students is the services they, and their families, receive from other government and nonprofit agencies including religious institutions, food banks, and social service agencies. To fully understand the resources available for each child, some knowledge of these services is also important. The most likely place to retrieve this kind of information is through the parent interview. McCroskey and Meezan (1997) show that there is a very high correlation between parent self-reports on social services received and social worker reports on family receipt of these services. Thus, it might be possible to develop reasonable data on what other services the children received through the addition of items to the parent survey.

The ECLS-K provides an opportunity to examine measures of program costs for a representative group of school children as they move from kindergarten to the fifth grade.

In addition to public or quasi-public services, the time parents spend helping their children with homework after school is an important educational resource, as is knowledge of the parents' income and educational attainment. In addition, some measure of the number of books in the home, and whether or not the child's family has a computer in the home may provide information on resources available to each child that might help in linking educational resources to student outcomes—even if those resources are found outside the traditional school.

Conceptual Framework

Fit with ECLS-K

The Study Design Report for ECLS-K (Ingels et al. 1996) presents the conceptual framework for the ECLS-K. Resource indicators described in this section of the article fit within the "School and Classroom Characteristics, Policies and Practices" component of that conceptual framework. However, our recommendations add to the conceptual framework by including the issue of "school resources" to the framework.

A second issue has to do with resources available to a child and his or her family outside of the school. Ingles et al. include a "Community and Neighborhood Characteristics" component in their framework. The cost of these characteristics was added to this component of the framework. The question that needed to be resolved was whether or not collection of information that would allow for the estimation of costs could be done in a manner that protected the privacy of families.

Cost Model

The problem to be addressed when constructing a student-level questionnaire is how to collect cost data that allows one to develop a student resource measure. Berne and Stiefel (1997) describe cost concepts that help describe the best way to ascertain costs at the student level. They argue that there are two major approaches for measuring costs—departmental costing and product costing. Departmental costing estimates the costs of different administrative units, whereas product costing is designed to determine the cost of producing individual products. Berne and Stiefel (1997) argue that for development of a student resource measure, product costing is more relevant. Specifically, they state:

> In our study, the concepts of product costing are relevant because the questions to be answered with the resource data are centered, for the most part, on the student (product) and not on the administration of the districts or schools (units) that 'produce the education' for students. We are interested in whether changing the way resources are allocated to students will change outcomes for the students, whether resource allocations to students are equitable, and whether students are receiving the resources

BEST COPY AVAILABLE

> that are intended for them (Berne and Stiefel 1997, 73).

The task we faced was to identify the full costs of the product or in this case the full costs of providing education services to each student. Berne and Stiefel argue that determination of the full costs of providing education services needs to take both direct and indirect costs as well as both variable and fixed costs, into account. The model presented in this paper attempts to consider each of these types of costs, relying on a combination of job-order costing and process costing techniques.

Job-order costing determines the costs of each individual unit of a product, such as custom-made materials and equipment, whereas process costing determines the costs of groups of identical units and divides by the number of units to obtain an average cost (Berne and Stiefel 1997). The latter is most frequently used in analyses of educational expenditures. The ECLS-K, by focusing on individual students, offers an opportunity to, in some instances, identify individual student costs, allowing development of a hybrid model that will estimate student resources as accurately as possible.

The model presented in this paper attempts to consider each of these types of costs, relying on a combination of job-order costing and process costing techniques.

Ideally, all data necessary to develop a student resource measure would be collected through one ECLS-K survey instrument. Unfortunately, given the study design, this would not be possible. In fact, the single biggest expenditure item, the teacher salary, could not be collected accurately through the originally constructed surveys. Therefore, a School Business Administrator Questionnaire needed to be developed. This new questionnaire collects salary and benefit information on the teachers, aides, principal and counselors aiding a specified child.

Using the cost model suggested by Berne and Stiefel (1997), table 2 summarizes the items that need to be included to get an accurate measure of individual student resources in kindergarten, and the proposed approach for ascertaining the costs of each element. A more detailed discussion of each follows below. The goal of these additions was to ascertain the level of resources available for every student. Consequently, to the maximum extent possible, the data collection effort was designed to provide enough information to allow researchers to allocate student-specific costs directly to those individual students. The discussion below indicates whether the original ECLS-K questions provided adequate information to develop a student resource measure and how they could be used. In instances where current ECLS-K items were not adequate, additions to the questionnaires are provided.

Measures of each element identified in table 2 were not available in the originally planned ECLS-K instruments. Other than the School Business Administrator Questionnaire, it is unlikely that additional survey instruments can be developed and implemented at this time. As a result, the recommended approach was, where possible, to collect direct costs. In cases in which that approach was either too difficult, or potentially too expensive, then methods to estimate costs indirectly were suggested. The goal was to use the direct costing approach for as many of the costs as possible, and as decisions are made, to rely on direct approaches for expenditure items that represent larger proportions of total costs and rely on indirect costing methods for those representing smaller proportions of the total.

Approach

In this section, each cost item identified in table 2 is discussed in more detail, and the approach for determining student costs is described. In addition, ECLS-K items that relate to the cost item are identified and their potential use is discussed. Examples of the cost items include classroom costs, teacher compensation and instructional materials. In cases in which additional data are needed, additions to specific ECLS-K questionnaires are provided. An additional questionnaire focused on the district's chief business officer is also recommended. A copy of the final questionnaire is included in the appendix.

Classroom Costs

Teacher Compensation

Teacher compensation (salary and benefits) is the single largest component of spending on all students. Consequently, it is important to get as accurate a picture of

BEST COPY AVAILABLE

Table 2.—Cost approach recommended for Early Childhood Longitudinal Study (ECLS-K) student resource measure: Analysis of individual expenditure items

Variable	Direct cost approach	Indirect cost approach
1. Classroom level costs		
a. Teacher compensation		
i. Salary	X	
ii. Benefits	X	X
b. Instructional aide compensation		
i. Salary	X	
ii. Benefits	X	X
c. Instructional materials	X	
d. Special programs		
i. Special Education	X	X
ii. Com	X	X
iii. Gifted and Talented	X	X
iv. LEP programs	X	X
2. School level costs		
a. Site administration		X
b. Instructional support		X
c. Student support		X
d. Maintenance and operations		X
e. Utilities		X
f. Transportation		X
3. District level costs		
a. District administration		X
b. Facilities		X
c. Data processing		X
4. Non-school costs		
a. Other agency expenditures		X
b. Parental support		X

SOURCE: Author's sketch.

teacher salary as possible. A number of items in the ECLS-K questionnaires provide information about factors that impact individual teacher salary, but none specifically seeks the salary of an individual teacher. There appear to be three options for getting this information:

1. Ask the teacher. This is unreliable and will not include the costs of benefits.
2. Ask the principal. The principal may not know individual teacher salaries, certainly will not know the exact benefit costs associated with individual teachers, and may not know the benefit rate applied to teacher salaries in the district.
3. Ask the school business office. This is potentially the most reliable way to get accurate teacher costs, and it would need to be done during site visits and follow-up. It could be problematic as it is a new source of data for the survey, further complicating an already complicated field visit task. Moreover, it would require the data collectors to contact someone in the district's business office and get information that is often treated as confidential. A survey of eight school business officers in districts in eight states indicated that all of them could provide accurate salary information for individual teachers. Seven of the eight indicated that it is also possible to provide accurate benefit cost data for individual employees, while the eighth indicated that it could only provide benefit cost data as a percentage of salary for classified or certificated employees. All but one of the districts indicated that the information is public, and if

the site data collector asked, they would provide it. One of the business officers indicated that she could not give out information attached to an individual name, but said it would be easy to provide data on all of the kindergarten teachers at a school, (which is fine in year one since all kindergarten teachers will be surveyed), and that it could probably even be provided by room number meaning that it would be possible for the on-site data collector to attach salaries to individual teachers and individual students which is the goal of the collection of student resource information.

It was recommended that a short questionnaire be developed and administered to the chief business office of each school district. It could best be done via an in-person or telephone interview. However, much of the data sought will probably require some research on the part of administrators or their staff. The final form of the School Business Administrator Questionnaire appears in the appendix. The survey seeks four specific pieces of information for each kindergarten teacher: the base salary (from the district's salary schedule), any supplemental pay that individual receives, the benefit rate the district applies to that salary, and whether the individual worked part- or full-time. The questionnaire also requests salary information for school principals.

Instructional Aide Compensation

To the extent they are used in kindergarten classes, instructional aides represent a substantial expenditure on behalf of each student. The same issues raised for collection of teacher compensation apply here as well. The questionnaire for school business administrators was designed to collect information on aides as well as teachers.

Instructional Materials

Another major component of resources available to individual students are the instructional materials in a classroom. A number of items on the questionnaires provide information on the types of instructional materials available in each of the kindergarten classrooms in the sample. Table 3 summarizes the relevant items, their location in ECLS-K questionnaires, and indicates the types of instructional materials each item seeks to identify.

Table 3 shows that information is available regarding most of the instructional materials available for kindergartners. Missing is some assessment of the number of books available in the classroom. The spring Kindergarten Teacher Questionnaire (part A) asks about the adequacy of textbooks and tradebooks. Unfortunately, the question elicits a subjective response from teachers that does not allow the analyst to ascertain how many books or their estimated value that could be included in a resource measure. Moreover, if textbooks are not adequate, it is not clear from the answer why they are inadequate. Is it because of the number of books, age, or their condition?

Instructional materials represent a relatively small amount in the overall budget of a school. It might be best to simply leave the questions as they are, which would enable researchers to analyze and compare adequacy of books, computers, and other equipment across classrooms. Another approach would be to ask a series of questions of each teacher about their access to these materials. If the latter approach is chosen, the following question could be added to the spring Kindergarten Teacher Questionnaire (part A).

X1. How many different textbooks do you use in your class? _____

Table 3.—Items identifying instructional materials in Early Childhood Longitudinal Study (ECLS-K) questionnaires

Materials identified	ECLS questionnaire	Item number
Curriculum materials	Special Education Teacher	15
Number of computers for institution and administration in school	School Principal	27
Rooms with computers and capability of computers	School Principal	28
Adequacy of various instructional materials in classroom	Kindergarten Teacher (spring) part A	26
Availability of different types of activity areas in classroom	Kindergarten Teacher (fall) part B	1
Access to a home computer and its use	Parent Interview spring	HEQ220, HEQ230 (p.58)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K).

Are there enough of each textbook for each enrolled child?

_____ Yes (skip to X2)

_____ No

If no, approximately what percentage of the children in your class have copies of each textbook used?

Book 1 _____ %

Book 2 _____ %

Book 3 _____ %

Book 4 _____ %

Book 5 _____ %

X2. For each of the following, indicate whether you have the item in your classroom, or if you share with other classes. (For each item place an X in the appropriate box and if it is shared indicate how many other classrooms also use that item.)

Place an X in one Box only

Item	Have	Don't have	Share	Number of classrooms sharing item
Television				1 2 3 4 5 or more
VCR				1 2 3 4 5 or more
Manipulatives (e.g. Blocks and puzzles)				1 2 3 4 5 or more
Video tapes and film				1 2 3 4 5 or more
Film projector				1 2 3 4 5 or more
Ditto or photocopier equipment				1 2 3 4 5 or more
Art materials				1 2 3 4 5 or more
Musical instruments				1 2 3 4 5 or more
Musical recordings				1 2 3 4 5 or more
Materials for teaching LEP children				1 2 3 4 5 or more
Materials for teaching children with disabilities				1 2 3 4 5 or more
Computers				1 2 3 4 5 or more
Computer software				1 2 3 4 5 or more

X3. How many computers are located in your classroom? (circle the number)

1 2 3 4 5 6 7 8 9 10 or more

Another element that might be useful in understanding resources available for students is the quality of the school library and the resources available there. Probably the best way to retrieve this information is to add the following questions to the Principal Questionnaire.

X4. Approximately how many books are there in your school library/media center? _____

What percentage would you estimate are appropriate for kindergarten? _____%

X5. What other instructional materials appropriate for kindergarten are available in your media center (check all that apply)

_____ Computers
_____ Film
_____ Video tapes
_____ Audio tapes
_____ Other (please identify)
_____
_____
_____

Special Programs

Crucial to accurately measuring the full range of resources available to students is an understanding of the special program services each student receives. Among the most important are special education, compensatory education, gifted and talented education, and LEP programs. Table 4 identifies the questionnaire items that provide information on special programs for kindergarten students. Each of these four areas is addressed separately below. It should be pointed out that for all four areas, item 23 on the Principal Questionnaire seeks information on sources of funding, including programs covering each of these areas.

1. **Special Education:** Potentially the most expensive of these programs, it is important to know what resources are available to each child who is identified as having a learning disability. While it is unlikely that many kindergarten students will be identified with more moderate disabilities, many with more severe disabilities will certainly require services by the time they arrive in kindergarten (and in many cases will already be receiving services from the school district). Since the sample includes children in special day classes (full-time special education), additional efforts are not required to identify these programs and the resources they receive. Questions 1, 3, and 4 in Kindergarten Teacher Questionnaire (spring) part A can be used to identify classes that are for special education children. Those items ask for the number of children in the class, and the number with various types of disabilities. Should the number of children correspond with the number with disabilities, it is likely that the class is designed for children with disabilities.

The problem arises for children with disabilities who are either included in a regular education class, or receive some type of pull-out instruction. It is these children for whom additional information may be required. If researchers are to estimate the resources available for them, they must include the costs of additional equipment in the regular classroom, or an estimation of the costs of the pull-out program.

It is beyond the scope of this study to collect detailed data on individual special education classes. However, if it were possible to know how many hours per day or week a child is removed from the regular classroom for special instruction, a researcher using estimates of the costs of different special education programs would be able to estimate the share devoted to that child and offer a reasonable estimate of the additional resources devoted to the child due to his or her disability.

This can be determined relatively easily through the Student Information Checklist. For special education, two items could be added. The first item would ask if the child receives services outside of the regular classroom and the next would ask for how many hours. The specific wording would be:

Table 4.—Items related to special program enrollment and services in Early Childhood Longitudinal Study (ECLS-K) questionnaires

Special program or service	ECLS questionnaire	Item number
Student disability	Special Education Teacher	1
Hours per week of direct special education	Special Education Teacher	9
Services provided to disabled student	Special Education Teacher	10
Primary placement of disabled student in general education	Special Education Teacher	12
Hours spent in special education per week	Special Education Teacher	13
Teaching practices and methods used with student	Special Education Teacher	14
Evaluations received by student	Special Education Teacher	23
Does student have an IEP	Student Records Abstraction Form	7
Is student LEP	Oral Language Development Scale Subtest Form Pre-LAS 2000	2e
Type of ESL instruction	Student Information Checklist	2e, f
Participation in Gifted and Talented Program	Student Information Checklist	2g
Receipt of individual tutoring	Student Information Checklist	2a, c
Enrolled in program for emotional or behavioral problems	Student Information Checklist	2h
Number of years a child receives ESL, bilingual instruction or both	School Principal	51
Number of children in special programs by type of program	Kindergarten Teacher (spring) part A	5
Number of LEP students, type of instruction received, time spent on language instruction in classes	Kindergarten Teacher (fall) part A	16, 17, 19
Number of children with disabilities in class	Kindergarten Teacher (spring) part A	3
Number of children who need special services	Kindergarten Teacher (spring) part A	5
Number receiving various special education services	Kindergarten Teacher (spring) part A	5

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K).

	Student 01	...	Student 15
1x Receives special services outside of the regular classroom	Y	Y	Y
1y Number of hours per week that services are provided	N	N	N

2. Compensatory Education: This category of services receives the least attention in the ECLS surveys. It is recommended that the same strategy used to estimate the resources available to students eligible for special education be used for contemporary education. This would lead to the addition of the following items in the Student Information Checklist:

	Student 01	...	Student 15
1x Receives services through Title I or a state compensatory program	Y N	Y N	Y N
1y Receives In-class or Pull-out compensatory services (circle I or P)	I P	I P	I P

This item does not ask for the number of hours per week that pull-out programs are provided. It is assumed that cost estimates could be derived from the many studies of Title I that have been conducted in the past and are being considered for the future.

3. LEP Programs: Items 1d and 1e in the Student Information Checklist collect data on the LEP status of each child in the sample and the type of services the child receives. While data on the costs of various bilingual programs is less available than data on the costs of special and compensatory education programs, it would still be possible with this information to estimate the additional resources available to each LEP child. It seems that no additional data need to be collected for this particular category of programs.

4. **Gifted and Talented Programs:** This is the most problematic of the four programs considered. Gifted and Talented programs are not as common as the others identified above, and are often funded individually by districts or even schools. There is less information on what types of services are offered to students, and the literature on the costs of these programs is almost non-existent. Moreover, it is unlikely that school or district officials will have accurate data on these programs. It is also hard to imagine that at the kindergarten level, such distinctions are made quite frequently. Perhaps one of the most important finding from the ECLS-K will be to give researchers a better sense of the number of Gifted and Talented programs available for children at this young age. While specific cost and resource data will be unavailable at this time, it seems unlikely to dramatically impact the estimation of student-level resources. For these reasons, it seems that at the present time, there is no need to attempt to collect additional information on Gifted and Talented programs through the ECLS-K.

School-Level Costs

Table 2 identifies six different school-level costs: site administration; instructional support; student support; maintenance and operations; utilities; and transportation. With the exception of utilities and transportation, the bulk of these costs are also related to personnel. Item 58 in the Principal Questionnaire seeks information on the full-time equivalent (FTE) number of staff in a wide range of categories which would make it possible to allocate portions of the costs of each of these services to kindergarten students. There is one change that would be helpful. Item 58r asks for the number of "all other noninstructional staff (include maintenance, food service, and clerical staff)." It would be helpful if those staff components were broken out separately. Specifically, item 58 should be expanded to add the following staff classifications after item 58r:

s. Office clerical staff

t. Custodial staff

u. Food service staff

These additions would make it possible to estimate the costs of personnel for site administration when combined with the average salary data provided by the district's business administrator. Missing is an estimate of the costs of supplies and materials for the office and for the custodial staff. These amounts represent a very small portion of the total budget of a school and it is unlikely respondents to the ECLS-K survey will have accurate information, unless the school has a detailed site budget. While

one could ask for this information, it seems best to let researchers rely on estimation techniques based on other readily available school expenditure data.

This leaves estimates of the cost of utilities and transportation. Utility costs are probably best identified by the school business administrator and applied uniformly across the district. While this is not completely accurate since some schools will be more energy efficient than others, a per student average for the district will provide most researchers with adequate information to distribute these costs among students in a school. The following question could be included in the School Business Administrator Questionnaire:

> What would you estimate are the per pupil utility costs incurred by your district this year? $_____ per pupil.

Transportation is more difficult. It is only through the Parent Questionnaire, specifically item P1Q310 on the spring parent kindergarten survey, that the way a child gets to school each day is identified. The responses to this question could be used to count the number of children who are bussed to school. With cost estimates from the district business administrator, an estimate could be added to the student resource measure. The difficulty of this approach is that transportation costs are typically maintained in terms of cost per mile or cost per hour of operation. If costs per mile could be provided, it might be possible to add an item to P1Q310 on the parent questionnaire seeking the approximate distance a child must travel to get to school. Specifically, the following would be added after question P1Q310:

> Approximately how far in miles is your home from the school CHILD attends? _____

In addition, the following question would have to be added to the School Business Administrator's Questionnaire:

> Approximately how much do you spend per mile for pupil transportation? $_____ per mile

District-Level Costs

Previous research (see for example, Picus and Fazal 1996) has shown that central office administration represents a relatively small portion of total school district expenditures. The most straightforward approach to ascertaining these costs for inclusion in the student resource measure is to ask the district's business administrator:

> What are your estimated per pupil costs for central administration for the current year? $_____ per pupil.

This figure will probably provide an excellent estimate for the student-level analysis.

Non-School Costs

Estimating the costs of non-school programs available to each child is the most complex part of the analysis, and may not be important to all researchers. At this point, rather than try to estimate the costs of these services, it seems prudent to be satisfied with the data that will be collected from the Parent Questionnaires providing information on the services that children and their families receive. The questionnaires collect data on child care, Head Start, AFDC, food stamps, and many other services. Simply being able to list all of them for the student sample will be a major improvement over current school-level data.

Recommendations

The discussion above summarizes the recommendations that were made to NCES regarding the collection of student-level fiscal data through the ECLS-K survey instruments. As shown in the discussion above, the recommendation was to add approximately 14 new questions in total to the already existing questionnaires and checklist while also adding a new School Business Officer Questionnaire. The primary purpose for the additional questions was to elicit further information regarding the resources available to the identified student, and to identify programs in which the child participates. Examples of additional resource questions include questions to the principal regarding the number of books in the school library and instructional materials in the media center (e.g., audio and video tapes, film, computers, etc.). Also requested was information from the teacher as to the range and number of textbooks used in the class and the number of computers located in the classroom. To ascertain information about program involvement, additional questions were recommended for the Student Information Checklist including whether the child receives special services outside of the regular classroom. If the answer is yes, then additional information is requested such as the number of hours per week that the service is provided,

whether it is funded through a federal program and if it is an in-class or pull-out program. The purpose of all of these questions is to gain a more refined understanding of the actual resources the student being studied receives. In addition to these questions, it was recommended that a School Business Officer Questionnaire be developed to collect information on teacher base salary, merit pay, benefit expenditures, and an indication of whether or not the individual was a full- or part-time employee.

Arguably, the additional information recommended for collection combined with the salary, merit pay, and benefit data to be collected through the proposed new School Business Officer Questionnaire and the existing ECLS-K questionnaire items would have served as the most comprehensive and rich data set on the education of children to date. This information would have helped researchers to begin to discern which educational programs and resources provide the greatest benefits for their costs and what optimal educational programs look like.

NCES Implementation

Unfortunately, due to the balance that needs to be struck between maximizing the substantive areas covered by a questionnaire and minimizing respondent burden, the inclusion of all our recommendations in the ECLS-K instruments was not possible. Instead, NCES agreed to add the School Business Administrator Questionnaire (for public schools with a similar questionnaire developed and administered to Catholic and other private school principals) to ascertain the base salary, merit pay, and employee benefits for the principal and teachers serving the ECLS-K sampled students. Appendix A includes copies of this questionnaire form.

Additionally, NCES developed and included a second new instrument, the ECLS-K Facilities Checklist. The new Facilities Checklist was designed to measure the physical environment and social climate of kindergartner's schools. The Checklist was completed by field interviewers based on their observations of various aspects of students' school facility and student interactions in school. Appendix B includes a copy of the ECLS-K Facilities Checklist.

Value Added

Clearly, we were not able to do all that we had wished, but we have made important advancements. A critical question to be answered now though is what will we get for this effort? Berne and Stiefel (1997) argue "it is crucial that we begin to make progress on the question of effective use of resources in education." A database with resource measures linked to student outcome measures is critical to understanding how resources can be used effectively.

The value of understanding how resources are linked to student outcomes lies partially in helping school districts make wise resource allocation decisions.

Since ECLS-K will assess the performance of a large sample of children for a period of six years, it is theoretically possible to understand how resources and student outcomes are linked (assuming such a linkage exists). The value of understanding how resources are linked to student outcomes lies partially in helping school districts make wise resource allocation decisions. Even if the information that is collected is limited and the expenditure data imperfect, the insight into the services available to each child and costs associated with those services will be helpful in understanding how and why resources matter.

BEST COPY AVAILABLE

References

Berne, R. and Stiefel, L. May 1997. "Student-Level School Resource Measures." In Fowler, W. J. (ed). *Selected Papers in School Finance 1995.* Washington, DC: National Center for Education Statistics.

Ferguson, R. F. 1991. "Paying for Public Education: New Evidence on How and Why Money Matters." *Harvard Journal on Legislation,* 28, 465–497.

GAO. 1996. *School Facilities: America's Schools Report Differing Conditions.* Washington, DC: General Accounting Office. GAO/HEHS-96-103.

Greenwald, R., Hedges, L. V., and Laine, R. D. 1996. "The Effect of School Resources on Stuent Achievement." *Review of Educational Research,* 66(3), 361–396.

Hanushek, E. A. 1989. "The Impact of Differential Expenditures on School Performance." *Educational Researcher, 18(4),* 45–51.

Hanushek, E. A. 1994a. "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald." *Educational Researcher,* 23(4), 5–8.

Hanushek, E. A. 1994b. *Making Schools Work: Improving Performance and Controlling Costs.* Washington, DC: The Brookings Institution.

Hanushek, E. A. 1996a. "The Quest for Equalized Mediocrity: School Finance Reform Without Consideration of School Performance." In Picus, L. O. and Wattenbarger, J. L. (eds). *Where Does the Money Go? Resource Allocation in Elementary and Secondary Schools.* Thousand Oaks, CA: Corwin Press. 20–43.

Hanushek, E. A. 1996b. "A More Complete Picture of School Resource Policies." *Review of Educational Research,* 66(3), 397–410.

Hedges, L. V., Laine, R. D. and Greenwald, R. 1994a. "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes (An exchange part 1)." *Educational Researcher,* 23(3), 5–14.

Hedges, L. V., Laine, R. D. and Greenwald, R. 1994b. "Money Does Matter Somewhere: A Reply to Hanushek." *Educational Researcher,* 23(4), 9–10.

Hertert, L. 1996. "Does Equal Funding for Districts Mean Equal Funding for Classroom Students? Evidence From California" in Picus, Lawrence O. and Wattenbarger, J. L. eds. *Where does the Money Go? Resource Allocation in Elementary and Secondary Schools.* 1995 Yearbook of the American Education Finance Association. Newbury Park, CA: Corwin Press.

Ingels, S. J., S. L. Dauber, T. B. Hoffer, J. Mulrow, L. A. Scott, and J. R. Taylor. 1996. "Early Childhood Longitudinal Study, Kindergarten Cohort: Study Design Report." Chicago: NORC, report to National Center for Education Statistics.

Ladd, H. and Ferguson, R. 1996. Ladd, H. F. (ed.) *Holding Schools Accountable.* Washington, DC: The Brookings Institution. 299–326.

Laine, R. D., Greenwald, R. and Hedges, L. V. 1996. "Money Does Matter: A Research Synthesis of a New Universe of Education Production Function Studies." In Picus, L. O. and Wattenbarger, J. L. (eds). *Where Does the Money Go? Resource Allocation in Elementary and Secondary Schools.* Thousand Oaks, CA: Corwin Press. 44–70.

Levin, H. 1983. *Cost Effectiveness: A Primer.* Beverly Hills, CA: Sage Publications.

McCroskey, J. and Meezan, W. 1997. *Family Preservation and Family Functioning.* Washington, DC: Child Welfare League of America.

Nakib, Y. A. 1996. "Beyond District-Level Expenditures: Schooling Resource Allocation and Use in Florida." In Picus, Lawrence O. and Wattenbarger, J. L. eds. *Where does the Money Go? Resource Allocation in Elementary and Secondary Schools.* 1995 Yearbook of the American Education Finance Association. Newbury Park, CA: Corwin Press.

Picus, Lawrence O. and Fazal, Minaz B. 1996. "Why Do We Need to Know What Money Buys? Research on Resource Allocation Patterns in Elementary and Secondary Schools." In Picus, Lawrence O. and Wattenbarger, J.L. eds. *Where does the Money Go? Resource Allocation in Elementary and Secondary Schools.* 1995 Yearbook of the American Education Finance Association. Newbury Park, CA: Corwin Press.

Picus, Lawrence O. 1993a. *The Allocation and Use of Educational Resources: School Level Evidence from the Schools and Staffing Survey.* Los Angeles, CA: USC Center for Research in Education Finance. Working Paper No. 37.

Picus, Lawrence O. 1993b. *The Allocation and Use of Educational Resources: District Level Evidence from the Schools and Staffing Survey.* Los Angeles, CA: USC Center for Research in Education Finance. Working Paper No. 34.

Picus, Lawrence O. 1997. "Does Money Matter in Education? A Policymaker's Guide." In Fowler, W. (ed). *Selected Papers in School Finance, 1995.* Washington, DC: National Center for Education Statistics.

U.S. Department of Education. National Center for Education Statistics. 1999. *Digest of Education Statistics, 1998.* Washington, DC: NCES 1999–036.

Appendix A. School Business Administrator Questionnaires

Early Childhood Longitudinal Study-Kindergarten Cohort
ECLS-K

Catholic School Principal
Salary and Benefits Questionnaire

In order to trace resources directly available to the children in our sample, please provide the base salary for the individuals listed below.

- Base Salary: The gross salary earned by the individual. Please include any additional funds received for having a Masters (MA) or Ph.D.
- Merit Pay: Includes any additional stipends the educator receives for exemplary work. (Please do not include the base salary in this figure).
- Employee Benefits: Includes any payroll taxes, retirement, medical, dental, disability, unemployment, life insurance, and other fringe benefits (e.g., unused sick leave). (Please do not include base salary in this figure). Please only include benefits that are employer paid.

Please also indicate with an "X" whether the professional is a full- or part-time employee.

School Name: «Sch_NAME» ID Number: «Sch_ID»

	Base Salary	Merit Pay	Employee Benefits	Full-Time	Part-Time
«Principal» «Prin_ID» (Principal)	$	$	$		
«Teacher1» «T1_ID»	$	$	$		

Thank you for completing this questionnaire. We appreciate your help and time. Please return the form in either the business envelope that was included in the packet or fax the form to:

Westat
Fax Number: (301) 963-5466

If you have any questions about this questionnaire, please call 1-800-750-6206.

Thank you again for your participation.

1

BEST COPY AVAILABLE

Early Childhood Longitudinal Study-Kindergarten Cohort
ECLS-K

Private School Principal
Salary and Benefits Questionnaire

In order to trace resources directly available to the children in our sample, please provide the base salary for the individuals listed below.

- Base Salary: The gross salary earned by the individual. Please include any additional funds received for having a Masters (MA) or Ph.D.
- Merit Pay: Includes any additional stipends the educator receives for exemplary work. (Please do not include the base salary in this figure).
- Employee Benefits: Includes any payroll taxes, retirement, medical, dental, disability, unemployment, life insurance, and other fringe benefits (e.g., unused sick leave). (Please do not include base salary in this figure). Please only include benefits that are employer paid.

Please also indicate with an "X" whether the professional is a full- or part-time employee.

School Name: «Sch_NAME» ID Number: «Sch_ID»

	Base Salary	Merit Pay	Employee Benefits	Full-Time	Part-Time
«PrinFirstName» «PrinLastName» (Principal)	$	$	$		
«Teacher1» «T1_ID»	$	$	$		

Thank you for completing this questionnaire. We appreciate your help and time. Please return the form in either the business envelope that was included in the packet or fax the form to:

Westat
Fax Number: (301) 963-5466

If you have any questions about this questionnaire, please call 1-800-750-6206.

Thank you again for your participation.

BEST COPY AVAILABLE

Early Childhood Longitudinal Study-Kindergarten Cohort
ECLS-K

Public School Business Administrator
Salary and Benefits Questionnaire

In order to trace resources directly available to the children in our sample, please provide the base salary for the individuals listed below.

- Base Salary: The gross salary earned by the individual. Please include any additional funds received for having a Masters (MA) or Ph.D.
- Merit Pay: Includes any additional stipends the educator receives for exemplary work. (Please do not include the base salary in this figure).
- Employee Benefits: Includes any payroll taxes, retirement, medical, dental, disability, unemployment, life insurance, and other fringe benefits (e.g., unused sick leave). (Please do not include base salary in this figure). Please only include benefits that are employer paid.

Please also indicate with an "X" whether the professional is a full- or part-time employee.

School Name: «Sch_NAME» ID Number: «Sch_ID»

	Base Salary	Merit Pay	Employee Benefits	Full-Time	Part-Time
«Principal» (Principal)	$	$	$		
«Teacher1» «T1_ID»	$	$	$		

Thank you for completing this questionnaire. We appreciate your help and time. Please return the form in either the business envelope that was included in the packet or fax the form to:

Westat
Fax Number: (301) 963-5466

If you have any questions about this questionnaire, please call 1-800-750-6206.

Thank you again for your participation.

Appendix B. ECLS-K Facilities Checklist

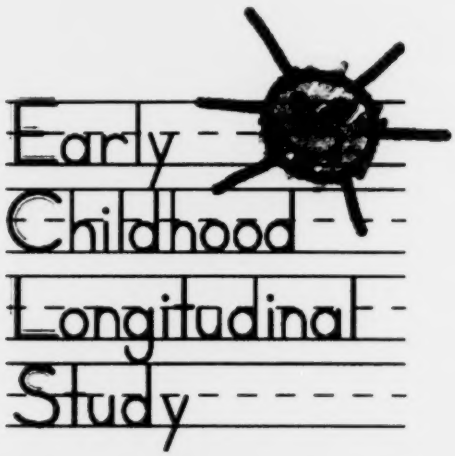
OMB #: 1850-0719
Expiration Date: 11/30/99

ECLS-K FACILITIES CHECKLIST

School Name: _____

ID #: _____ Date: _____

Field Supervisor's Name: _____ ID #: _____



Prepared for the U.S. Department of Education
National Center for Education Statistics by:

Westat
1650 Research Boulevard
Rockville, Maryland 20850

Assurance of Confidentiality

The collection of information in this survey is authorized by Public Law 100-297 and continued under the auspices of Section 404(a) of the National Education Statistics Act of 1994, Title IV of the Improving America's Schools Act of 1994, Public Law 103-382. Participation is voluntary. You may skip questions you do not wish to answer; however, we hope that you will answer as many questions as you can. No information collected under this authority may be used for any purpose other than the purpose for which it was supplied. Information will be protected from disclosure by federal statute (42 U.S. Code 242m, Section 308d). Data will be combined to produce statistical reports. No individual data that links your name, address, telephone number, or identification number with your responses will be reported.

BEST COPY AVAILABLE

ECLS-K FACILITIES RATING INSTRUMENT

1. For each of the following, circle "YES" if the facility was available to students; "NO" if the school does not have a facility or if it is not available to students. By available, we mean an area, separate from the classrooms, readily accessible to the students. For each of the facilities that are available and that you are able to observe, rate the condition and environmental factors using the following rating scale:

   (S) Satisfactory: The facility had enough space for purpose; bright light; open-air, airy; "like new" condition, minimal routine repairs needed, clean; comfortable, moderate noise.

   (U) Unsatisfactory: The facility was crowded; dim lighting; stale air; repairs required, dirty, graffiti; uncomfortable, stuffy (too hot or cold); loud distractive noise (e.g., screaming, yelling).

   The last column, Handicap Accessibility, refers to doors that are wide enough for wheel chairs, grab bars in bathrooms, and ramps and/or elevators in multi-floor buildings. Circle "YES" if these features are available; "NO" if these features are not available.

	Available	Observed	Space/ size	Light	Ventilation	Physical condition (ceiling, walls, floors, etc.)	Room temperature	Noise level	Handicap accessibility
a. Classroom			S U	S U	S U	S U	S U	S U	YES NO
b. Media center	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
c. Library	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
d. Art room	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
e. Music room	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
f. Cafeteria	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
g. Computer lab	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
h. Student bathrooms	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
i. Faculty bathrooms	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
j. Place for indoor play (gymnasium, multipurpose room)	YES NO	YES NO	S U	S U	S U	S U	S U	S U	YES NO
k. Outside playground	YES NO	YES NO	S U			S U			YES NO
l. School building (hallways, stairwells, common areas)		YES NO	S U	S U	S U	S U	S U	S U	YES NO

2. For each of the following, circle "YES" or "NO" for physical security in building. (CIRCLE ALL THAT YOU OBSERVED.)

	Yes	No	Yes, but not enforced
a. Security guard	1	2	3
b. Metal detectors	1	2	3
c. Security cameras	1	2	3
d. Window/door bars	1	2	3
e. Exit doors that only open from inside	1	2	3
f. Fencing around school	1	2	3
g. Sign-in policies	1	2	3
h. Visitors are greeted and directed by an adult to sign in at office	1	2	3
i. Internal communication system (e.g., intercoms)	1	2	3
j. Fire alarms	1	2	3
k. Fire extinguishers	1	2	3
l. Fire sprinklers	1	2	3

3. Based on the effectiveness of the security measures listed in question two, please rate your overall perceived feeling of safety for children in this school. Use the definitions provided below when making your choice. (CIRCLE ONE.)

Very Safe: The school has at least 6 of the safety measures listed in question two. All of these measures are in use and are effective. (If not all are in use, rate the school as "Safe".) No other safety measures are needed to protect the students.

Safe: The school has 4-5 of the security measures listed in question two. Although some additional measures could be added, the overall safety of the school is adequate. Most of the measures that the school does have are effective and in use.

Unsafe: The school has 2-3 of the safety measures listed in question two. Some of the measures are not enforced, and many more security measures are needed.

Very Unsafe: The school has less than 0-1 measures of security. Other security measures are definitely needed.

Very safe	Safe	Unsafe	Very unsafe
1	2	3	4

3a. Please indicate if the following factors are present in the neighborhood surrounding the school.

	A little	Some	A lot
a. Litter/trash	1	2	3
b. Graffiti	1	2	3
c. Boarded up buildings	1	2	3
d. Persons congregated on streets	1	2	3

4. Do you feel that the observed security is adequate? By observed security, we mean the measures listed in question two.

Yes	No
1	2

5. Below are some measures of happiness in schools. How many children did you observe doing the following? (CIRCLE ONE.)

Approximate number of children observed: _____

	None	A Few (2-4 children)	Many (5-10 children)	Most More than 10)
a. Fighting children	1	2	3	4
b. Laughing and/or smiling children	1	2	3	4
c. Crying children	1	2	3	4
d. Children talking/ chatting	1	2	3	4

6. Below are some measures of the overall learning environment in schools. Please tell us whether you agree or disagree that each measure was present in the school. (CIRCLE ONE.)

	Strongly agree	Somewhat agree	Somewhat disagree	Strongly disagree
a. Decorated hallways	1	2	3	4
b. Attentive teachers	1	2	3	4
c. Personable principal	1	2	3	4
d. Helpful staff	1	2	3	4
e. Order in hallways	1	2	3	4
f. Order in classrooms	1	2	3	4

United States
Department of Education
ED Pubs
8242–B Sandy Court
Jessup, MD 20794–1398

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

Standard Mail (B)

SPECIAL STANDARD MAIL

EP: 102303
ERIC CLEARINGHOUSE/ED/MGT
JAY STOECKL
U OF OREGON 1787 AGATE ST
EUGENE OR 97403